Crystal structure of the AAA+ and DNA-binding domain of the σ^{54}-activator ZraR

PhD thesis

László Sallai

European Molecular Biology Laboratory

Hamburg Outstation

2006

Part of this work was published in:

I. **Sallai L., Hendle J. and Tucker P. A.** X-ray crystallographic characterization and phasing of an NtrC homologue. ***Acta Crystallogr. D59***. 2003; 1656-1658.

II. **Sallai L. and Tucker P. A.** Crystal structure of the central and C-terminal domain of the sigma(54)-activator ZraR. ***J Struct. Biol.*** 2005; 151(2), 160-170

III. **Sallai L.** AAA ATP-áz fehérjék: egységes szerkezet, szerteágazó funkciók. ***Biokémia***. 2005; Dec.: 56-59.

Table of contents

Acknowledgements

Special thank to Dr. Paul A. Tucker for supervising and supporting this thesis. It was a pleasure working in the friendly and inspiring lab at the EMBL Hamburg Outstation.

I thank to the members of my thesis advisory committee, Dr. Matthias Wilmanns and Prof. Dr. László Dux, for support and helpful discussions.

Thanks to Dr. Jörg Hendle, who introduced me to the lab work; Dr. Peijian Zhou, who helped me with the molecular biology work; and Dr. Santosh Panjikar for giving me support at the computer.

Additionally, I would like to appreciate the former and current members of the lab; Alice, Attila, Carmen, Esben, Ingeborg, Ingrid, Katja, Nikos and Preben, for creating a very pleasant working atmosphere.

Introduction

1. The σ^{54}-dependent transcription

Transcription initiation by the *Escherichia coli* DNA-dependent RNA polymerase (RNAP) requires the association of the core RNAP ($\alpha_2\beta\beta`\omega$) with a sigma ($\sigma$) subunit. *E. coli* has seven σ factors with distinct functions that can be divided into two classes, the σ^{70} class and the σ^{54} class. The principal σ of *E. coli*, σ^{70}, is involved in the transcription of housekeeping genes as well as some non-essential genes and is the prototypical member of a large family of five related σ factors that are mainly involved in the response to various stresses. The *E. coli* σ^{54} protein belongs to a smaller family that is unrelated to the σ^{70} family in sequence and its main function is to regulate the nitrogen metabolism (Reitzer and Schneider, 2001).

The RNAP-σ^{54} complex initiates transcription by a mechanism quite distinct from RNAP containing the σ^{70}. Unlike the σ^{70}-RNAP, the σ^{54}-dependent promoters require distant activator sites that resemble enhancers in eukaryotic cells (Reitzer and Magasanik, 1986). These enhancer elements bind a distinct class of transcriptional activator proteins (enhancer binding proteins; EBPs) that hydrolyze ATP to increase the rate of open complex formation at the promoter (Fig. 1; Popham *et. al.*, 1989).

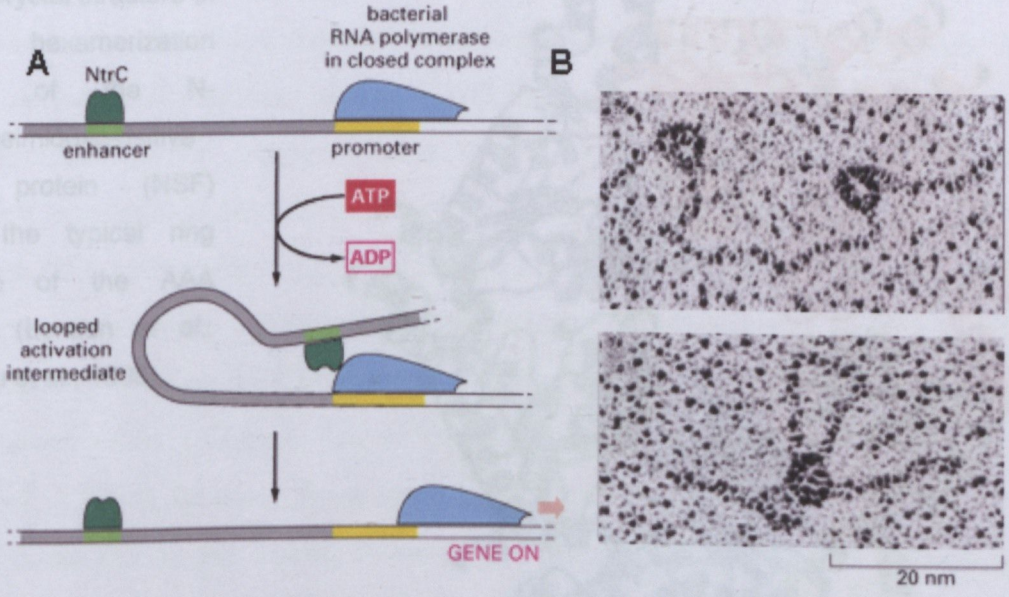


Fig. 1. (A) The general nitrogen regulatory protein NtrC, which belongs to the activators of the σ^{54} transcription factor, hydrolyses ATP and contacts the RNAP-σ^{54} complex by means of a DNA loop to initiate open complex formation. **(B)** Electron micrographs of NtrC and RNAP-σ^{54} on a DNA fragment containing the promoter-regulatory region of the glutamine synthetase gene in linear and looped conformation (Su *et al.*, 1990).

The interaction between the EBP and the RNAP-σ^{54} complex sometimes requires a DNA bending protein; the integration host factor IHF (Hoover *et. al.*, 1990) or the arginine repressor (Lu and Abdelal, 1999). The consequence and the biological advantage of these features is the strict control of gene expression; transcription of the σ^{54} dependent genes can be completely turned off. The most important control point of the σ^{54} dependent genes is the activator protein: phosphorylation, interaction with ligands or regulatory proteins modifies the ATPase activity (Shingler, 1996).

2. AAA+ proteins: Lords of the ring

Enhancer binding proteins that act upon the σ^{54}-containing form of the RNAP belong to the AAA superfamily of ATPases (ATPase associated with various cellular activities), which transform the chemical energy of ATP hydrolysis into mechanical force. Proteins in this family contain one or more copies of a 230-250 amino acid cassette that has ATPase activity (Patel and Latterich, 1998). The architectures of the proteins belonging to this superfamily are similar: the AAA ATPases form an oligomeric structure (Fig. 2.).

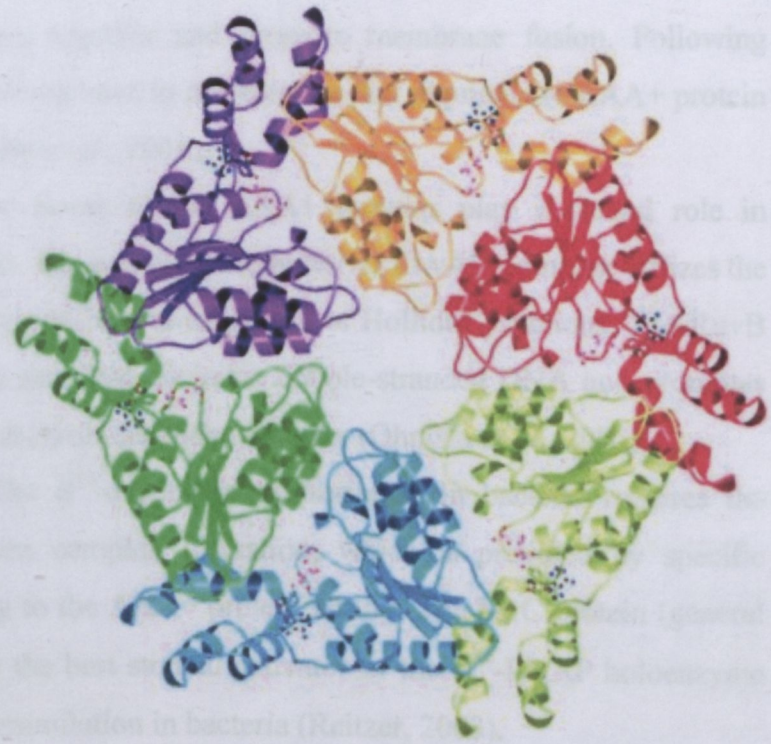


Fig. 2. Crystal structure of the hexamerization domain of the N-ethylmaleimide-sensitive fusion protein (NSF) shows the typical ring structure of the AAA proteins (Lenzen *et al.*, 1998, Yu *et al.*, 1998).

The AAA domain contains Walker A and B motifs, as well as several other motifs that distinguish it from classic P-loop NTPases. Motifs that are used to define the original AAA family have more recently been considered in combination with structural information, and this has led to further members being added to what is now referred to as the AAA+ family (Iyer *et al.*, 2004).

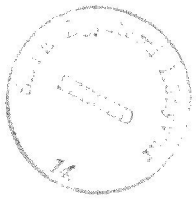
AAA+ proteins are found in many organisms and have diverse functions, from proteolysis to membrane trafficking. A simplifying idea that emerges from recent mechanistic studies is that AAA+ proteins all operate by promoting conformational changes or remodelling in target proteins. Functions of AAA+ proteins can be divided into four groups:

Protein unfolding and degradation: AAA+ proteins are crucial for proteolysis, in which their role is to unfold substrates and deliver them to the degradative chamber of so called chambered proteases (Pickart and Cohen, 2004). There are five AAA+ containing systems in bacteria (ClpXP, ClpAP, HslUV, FtsH and Lon) and equivalent systems in eukaryotes (the 26S proteasome and bacterial-like systems in organelles).

Promoting protein-complex disassembly: The second type of reaction catalysed by AAA+ proteins is the disassembly of stable protein complexes. One of the best studied systems is the SNARE (*S*oluble *N*SF *A*ttachment Protein *Re*ceptor), which assembles into complexes that bring membranes together and promote membrane fusion. Following membrane fusion, SNARE complexes have to disassemble that requires the AAA+ protein NSF to carry out this reaction (Jahn *et al.*, 2003).

Molecular motor proteins: Some of the AAA+ proteins play a crucial role in promoting macromolecule motion. For example, the bacterial RuvAB complex utilizes the energy of ATP hydrolysis to promote branch migration of Holliday junctions. The RuvB AAA+ protein hexamer forms a ring that encircles double-stranded DNA and promotes DNA translocation in an ATP hydrolysis-dependent manner (Ohnishi *et al.*, 2005).

Transcriptional factors: The σ^{54}-dependent transcription in bacteria requires the energy of ATP to promote open complex formation, which is provided by specific transcriptional factors that belong to the AAA+ protein family. The NtrC protein (general nitrogen regulatory protein C) is the best studied activator of the σ^{54}-RNAP holoenzyme complex that regulates nitrogen assimilation in bacteria (Reitzer, 2003).



3. Two-component signal transduction

Activators of the σ^{54}-RNA polymerase are often the regulator component of a two-component signal transduction system that can be characterised by phosphotransfer reactions between a sensor histidine kinase (HK) and a response regulator (RR). In prokaryotic cells the two-component signalling mechanism is the predominant signal transduction pathway (Stock *et al.,* 2000), for example, surveys of genomes have identified 62 two-component proteins in *E. coli* (Mizuno, 1997). Two-component phosphotransfer systems are present in fungi and plants as well (Loomis *et al.*, 1997; Stock *et al.*, 2000). So far the completed genomes of animals have not revealed two-component proteins, therefore they are potential candidates for antimicrobial therapy.

In typical two-component systems, sensor HKs monitor environmental stimuli and transmit this information to the RR by a phosphorylation event (Fig. 3). The HK core and the receiver domains are highly conserved, whereas the sensing domain is diverse. The autophosphorylation of the HK domains requires a dimer: one HK monomer phosphorylates the conserved His residue in the second monomer (Surette *et al.,* 1996). Unlike the typical protein kinase cascade in which one protein kinase phosphorylates multiple targets, in the two-component pathways the phosphoryl group is stoichiometrically transferred from the HK to the RR. Therefore, control in these systems is accomplished through the ability of the HK to regulate the phosphorylation state of the RR; many HKs have phosphatase activity to dephosphorylate their cognate RRs (Keener and Kustu, 1988; Aiba *et al.*, 1989; Lois *et al.*, 1993).

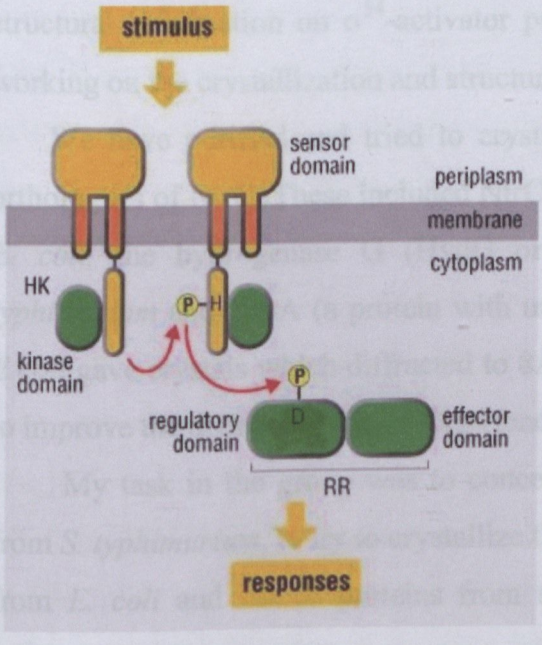


Fig. 3. Modular domains are involved in the two-component signal transduction. The basic two-component system utilizes a phosphotransfer pathway consisting of a single HK and RR. The HK contains a transmembrane sensor domain attached to an HK core consisting of a His-containing dimerization domain and an ATP-binding kinase domain. A single phosphoryl transfer occurs between a phospho-His of the HK and a phospho-Asp of the downstream RR, composed of an N-terminal regulatory (receiver) domain and an effector domain.

Two classes of two-component systems have been distinguished: the "orthodox" and the "hybrid" type (Stock *et al.*, 2000). The orthodox models of two-component systems predominate. They are composed of two proteins only, the HK and RR. Most orthodox HKs function as periplasmic membrane receptors, whereas the general nitrogen regulatory kinase (NtrB) is an example of a soluble cytoplasmic HK (MacFarlane and Merrick, 1985). The hybrid systems are characterized by a multistep phosphorelay that alternates between several histidine and aspartate residues (His-Asp-His-Asp) and involves additional receiver domains and a third type of conserved domains, the histidine-containing phosphotransfer module (HPt). The HPt domain contains a short consensus motif with an invariant histidine. The histidine typically obtains a phosphoryl group from a receiver domain and transfers it to another receiver domain within the signalling cascade. The signalling domains are either isolated or fused with each other in various combinations (Stock *et al.*, 2000). The osmoregulatory system of *Sacharomyces cerevisiae* is an example of the hybrid systems, in which the isolated HPt domain transfers the phosphoryl group to the effector protein and activates the mitogen-activated protein kinase (Maeda *et al.*, 1994). The orthodox systems are the most common architecture in prokaryotes, while phosphorelays, which provide a greater number of steps for regulation, are predominant in eukaryotes.

4. Aim

Though recent structural studies of AAA+ proteins provided insights into their general mechanical mode of action in using ATP binding and hydrolysis, the exact activation mechanism of the σ^{54}-dependent transcription remained unclear due to the lack of enough structural information on σ^{54}-activator proteins. To answer this question, we have been working on the crystallization and structural determination of AAA+ proteins.

We have purified and tried to crystallize a number of full length homologues and orthologues of NtrC. These included NtrC from *K. pneumoniae, Azotobacter vinlandii* and *E. coli,* the hydrogenase G (HydG or ZraR) protein from *E. coli* and *Salmonella typhimurium* and YfhA (a protein with unknown function) from *E. coli.* Only ZraR from *E.coli* gave crystals which diffracted to 8Å on a synchrotron beamline but we were unable to improve the diffraction limit (Jörg Hendle, personal communication).

My task in the group was to concentrate our efforts on shorter constructs of ZraR from *S. typhimurium,* to try to crystallize the phage shock protein A and F (PspA and PspF) from *E. coli* and –since proteins from thermophile organisms usually crystallize more

easily than the homologues in mesophiles– I was focusing on the NifA-like protein (AQ_218) from the hyperthermophile *Aquifex aeolicus* as well.

5 Target proteins

a. The Escherichia Coli phage shock protein A and F

The phage shock operon (*pspABCDE*) and the *pspG* gene of *E. coli* are specifically and continually induced by gene IV protein of filamentous phages. *Psp* genes are also synthesized in response to a variety of stress conditions or agents such as heat shock, osmotic shock and ethanol treatment (Brissette *et al.*, 1990, Lloyd *et al.*, 2004). The operon may play a significant role in enabling *E. coli* to compete for survival in the stationery phase under nutrient- or energy-limited conditions (Weiner and Model, 1994).

The product of the first gene, the PspA protein (MW. 25.5 kDa) is the most abundant in the cell upon induction (Brissette *et al.*, 1991). While PspA does not contain sequences characteristic of integral inner membrane proteins (Cserzo *et al.*, 1997), approximately 50% of the total cellular PspA is associated with the inner membrane of *E. coli*, and PspA is thus considered a peripheral membrane protein (Brissette *et al.*, 1990). The lack of any obvious DNA-binding motif suggests that PspA is not likely to bind DNA. Analysis of the protein sequence predicts four helices that form a coiled coil, a relatively unusual phenomenon in prokaryotes. PspA appears to maintain the proton motive force under stress conditions (Kleerebezem *et al.*, 1996) and can stimulate the export of secreted proteins (Kleerebezem and Tommassen, 1993). Cells lacking *pspA* have reduced viability under alkaline conditions as well as in late stationery phase (Weiner and Model, 1994). The three dimensional reconstruction of the PspA protein visualized by electron microscopy showed a 9-fold symmetric ring with an outer diameter of about 200 Å. Each of the nine domains in the ring is likely to be composed of four PspA proteins (Hankamer *et al.*, 2004).

Expression of the *pspABCDEG* genes requires their transcription by the σ^{54}-RNA polymerase and depends upon the PspF transcriptional activator protein (MW. 37 kDa), which is itself expressed independently of σ^{54} via a divergent σ^{70}-dependent promoter (Jovanovic and Model, 1997; Weiner *et al.*, 1995). PspF consists of a central AAA domain and a DNA-binding domain that are homologous to the corresponding NtrC domains (Fig. 4). The constitutively active protein differs from most members of the NtrC family in the absence of the N-terminal regulatory domain (Jovanovic *et al.*, 1996). More recently, in parallel with our work and the completion of this thesis, the three dimensional crystal structure of a PspF construct has been published that showed a typical ring-shaped AAA+

domain structure. The cryogenic electron microscopy resolution structure at 20 Å of PspF in complex with σ^{54} suggested the nucleotide-dependent conformational changes for association with σ^{54} (Rappas *et al.*, 2005; Rappas *et al.*, 2006).

The relationship between PspA and PspF and the regulation of the *psp* operon and the *pspG* gene lies in the negative control of PspA on the PspF activity. It has been shown both *in vitro* and *in vivo* that PspA inhibits transcription from the *psp* operon, which was independent of PspF binding to its UAS (Dworkin *et al.*, 2000). PspA directly and specifically acts upon and binds to the AAA+ domain of PspF (Elderkin *et al.*, 2002). Thus, PspA is considered the "escaped" N-terminal regulatory domain of PspF. Structural data indicated that a 12:36 PspF:PspA complex might form and is likely to be associated with the inhibition of the *psp* operon (Hankamer *et al.*, 2004; Rappas *et al.*, 2005; Rappas *et al*, 2006).

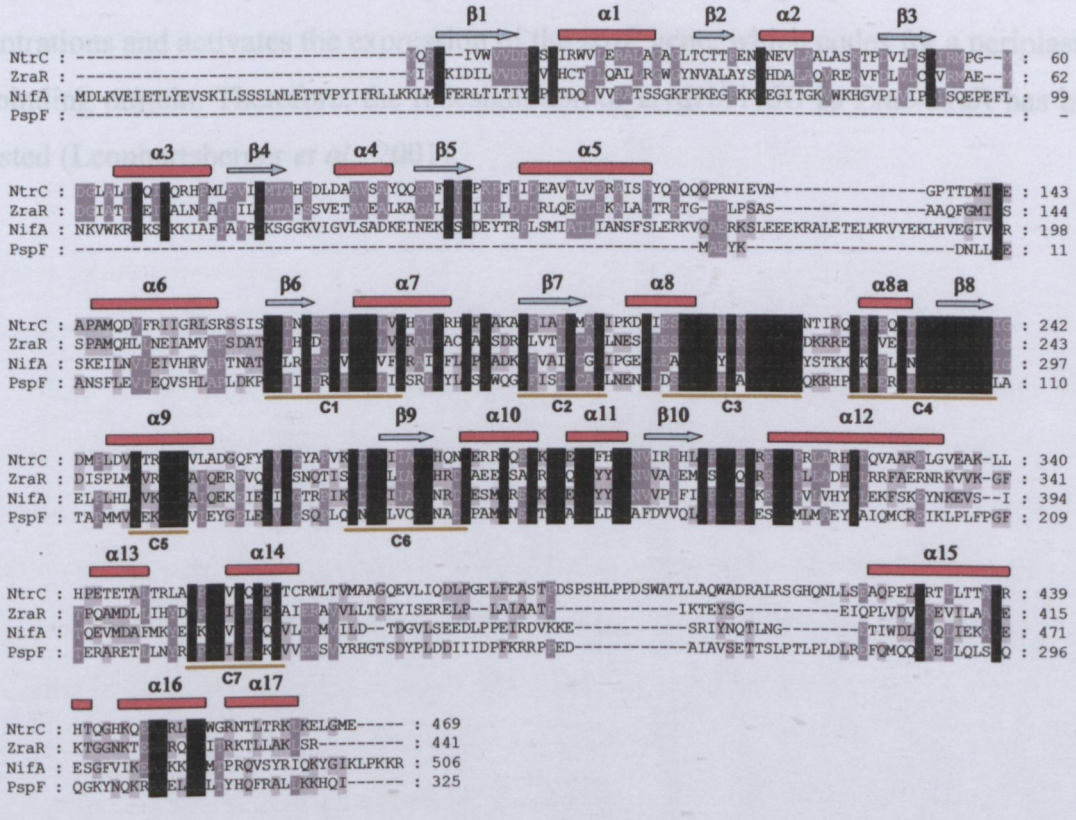


Fig. 4. Sequence alignment of NtrC and ZraR from *S. typhimurium*, NifA from *A. aeolicus* and PspF from *E. coli*. The secondary structure and the conserved regions (C1-C7) are indicated.

b. The NifA-like protein from Aquifex aeolicus

The *AQ_218* gene from the hyperthermophile *A. aeolicus* shows high sequence homology to those of the NtrC protein family (Fig. 4). Since the protein product of *AQ_218*, the NifA-like protein, binds with high affinity to the *A. aeolicus glnB* upstream region, it is presumably involved in the positive regulation of the *glnBA* operon and thus in the nitrogen assimilation (Studholme *et al.*, 2000).

c. The hydrogenase G protein from Salmonella typhimurium

E. coli and *S. typhimurium* can synthesise four different hydrogenases, designated hydrogenase 1, 2, 3 and 4. The protein products of the *hydH/hydG* genes have been postulated to play role in the regulation of the activity of hydrogenase 3 (Stoker *et al.*, 1989; Chopra *et al.*, 1991), which is part of the formate hydrogen lyase complex and is involved in the fermentative hydrogen production (Peck and Gest, 1957). It has been shown, that the HydH/HydG two-component system senses high periplasmic Zn^{2+} and Pb^{2+} concentrations and activates the expression of the *zraP* gene, which codes for a periplasmic Zn^{2+} binding protein. Therefore, the redesignation of HydH/HydG as ZraS/ZraR has been suggested (Leonhartsberger *et al.*, 2001).

Materials and Methods

1. Plasmids

Plasmid pET-28b (Novagen) containing the gene of the HTH deletion mutant of PspF (PspF-ΔHTH; res. 1-304); the *pspA* gene in pET-28b and the *AQ_218* gene in pET-15 (Novagen) were provided by Dr. Martin Buck (Dept. of Biological Sciences, Imperial College, London, UK). Plasmid pKH-420 with the gene of the central and C-terminal domains of ZraR (*zraR-CC*; res. 141-441) was a kind gift of Dr. Martin Drummond (Dept. of Molecular Microbiology, John Innes Centre, Norwich, UK).

2. DNA manipulations

A shorter construct of *pspF* containing the conserved central domain only (*pspF-C*; res. 1-253) was made and cloned into pETM-11 (pBR-322 modified by Dr. Günter Stier; EMBL Heidelberg, Germany). The *zraR-CC* gene was subcloned into pET-28a. The C146S mutation was introduced in the *pspF* genes using the Stratagene Quickchange mutagenesis kit. Plasmids were amplified and stored in either *E. coli DH-5α* or *XL-1Blue* cells. Plasmids were verified by sequencing in SEQLAB (Göttingen, Germany).

3. Protein overexpression

a. Overexpression of PspF-ΔHTH and PspF-ΔHTH C146S

E. coli B834 (DE3) pLysE cells were transformed with pET-28b::*pspF-ΔHTH* and pET-28b::*pspF-ΔHTH C146S*. The cells were plated on LB (Luria Bertani) agar with 50 mg/l kanamycin. A single colony was isolated and 2 L of LB media with antibiotic was inoculated. Cells were incubated at 37 °C till OD ~0.6 and induced with 0.5 mM of IPTG (isopropyl-beta-D-thiogalactopyranoside) for 3 h at 37 °C. Cells were centrifuged and resuspended in 50 mM NaH_2PO_4 pH=8.0, 500 mM NaCl and 5% glycerol.

b. Overexpression of PspF-C and PspF-C C146S

E. coli BL21 (DE3) cells were transformed with pETM-11::*pspF-C* and pETM-11::*pspF-C C146S* and grown on LB agar plates containing 50 mg/l kanamycin. A single colony was used to inoculate 2 L of LB with antibiotic. Cells were incubated at 37 °C until OD ~0.5 then the temperature was cooled down to 25 °C and the cells were induced with 0.5 mM IPTG. NaH_2PO_4 pH=8.0 was added to the medium to a final concentration of 20 mM and

the culture was incubated overnight. Cells were harvested by centrifugation and resuspended in 50 mM NaH_2PO_4 pH=7.0, 300 mM NaCl and 5% glycerol.

c. Overexpression of PspA

E. coli B834 (DE3) cells were transformed with pET-28b::*pspA* and grown overnight on LB agar plates containing 50 mg/l kanamycin. A single colony was inoculated into 1 l of LB media with the appropriate antibiotic and incubated at 37 °C until OD ~0.6. Cells were induced with 1 mM IPTG for 5 hours then harvested by centrifugation and resuspended in 100 mM Tris pH=7.5 and 300 mM NaCl. The suspension was frozen at -20 °C when it was not immediately used for protein purification experiments.

d. Overexpression of the NifA-like protein

E. coli BL21 Codon Plus cells were transformed with the pET-15::*AQ_218* plasmid and plated on LB agar containing 100 mg/l ampicillin. 2 L of LB media with ampicillin was inoculated and incubated at 37 °C till OD ~0.4. The media was cooled down to 25 °C and the culture was induced with 1 mM IPTG overnight. Cells were harvested by centrifugation, resuspended in 50 mM NaH_2PO_4 pH=7.0, 500 mM NaCl, 5% glycerol and frozen until use.

e. Overexpression of ZraR-CC

E. coli B834 (DE3) cells transformed with the pET-28a::*zraR-CC* plasmid were grown overnight at 37 °C on LB agar plates containing 50 mg/l kanamycin. A single colony was picked to inoculate 2 l of LB medium containing antibiotic. Protein induction was carried out at OD ~0.6 by the addition of IPTG to a final concentration of 0.5 mM. After induction the growth continued for 3 h at 37 °C. Cells were harvested by centrifugation and resuspended in 50 mM NaH_2PO_4 pH=8.0, 500 mM NaCl, 5% glycerol that was degassed to ensure the O_2-free environment. The suspension was either immediately used for protein purification or stored at -20 °C until use. The selenomethionine (Se-Met) containing protein was overexpressed in M9 media supplemented with Se-Met overnight at 37 °C.

4. Protein purification

a. Purification and thrombin cleavage of PspF-ΔHTH and PspF-ΔHTH C146S

Cells were broken up by sonication for 3x3 m and centrifuged at 20,000 g for 20 m. The supernatant was loaded on a Ni-NTA (nitrilotriacetic acid) column and eluted with the linear gradient of 50 mM NaH_2PO_4 pH=8.0, 500 mM NaCl, 5% glycerol and 1 M imidazole. The buffer was exchanged to 20 mM Tris pH=8.0, 50 mM NaCl, 2 mM DTT, 0.1 mM EDTA to which 20% glycerol was added and the sample was flash-frozen in liquid nitrogen when we did not perform crystallization experiments immediately. When the His-tag was removed, the sample was dialyzed against 20 mM Tris pH=8.0, 160 mM NaCl, 2 mM $CaCl_2$ and 2 mM DTT to which 15U/protein mg bovine thrombin was added and the sample was incubated at room temperature for 20 h. After thrombin cleavage, the sample was loaded on a Q-Sepharose column and eluted with the linear gradient of 20 mM Tris pH=8.0, 2 M NaCl and 2 mM DTT. Fractions containing the protein were dialyzed against 20 mM Tris pH=8.0, 50 mM NaCl, 2 mM DTT, 0.1 mM EDTA.

b. Purification of PspF-C and PspF-C C146S

The cell suspension was incubated on ice in the presence of lysozyme for 20 m which was followed by sonication for 3x1 m. The cell debris was harvested by centrifugation at 15,000 g for 20 m. The supernatant was loaded on a Ni-NTA column and eluted with the linear gradient of 50 mM NaH_2PO_4 pH=7.0, 300 mM NaCl, 5% glycerol and 1 M imidazole. Fractions containing the protein were loaded on a gel filtration column and eluted with the buffer of 10 mM Tris pH=8.5 and 150 mM NaCl. The protein solution was dialyzed against 10 mM Tris pH=8.0, 50 mM NaCl, 0.1 mM EDTA and 1 mM DTT which was supplemented with 20% glycerol and the sample was frozen when the protein was not crystallized immediately.

c. Purification of PspA

Purification of PspA was performed as described by Elderkin *et al.*: Cells were broken up by treatment with lysozyme and sonication for 3x1 m. The cell extract was centrifuged at 15,000 g for 20 m. The pellet was homogenized in 13.5 ml of 100 mM Tris pH=7.5, 300 mM NaCl, to which 2.6 ml of 3 M NaCl (final volume ~20 ml) and CHAPS (3-[(cholamidopropyl) dimethylammonio]-propanesulfonate) to a final concentration of 1.1 % were added to solubilize the protein. The suspension was rocked on a rotator platform for 2 h at 4 °C then centrifuged at 15,000 g for 20 m. The supernatant was added to 2 ml of Ni-

NTA resin and incubated on a rotator platform for 2 h at 4 °C. The resin was washed with 10 ml of 100 mM Tris pH=7.5, 300 mM NaCl, 60 mM imidazole and the protein was eluted with 1-1 ml fractions of 100 mM Tris pH=7.5, 300 mM NaCl and 1 M imidazole. 80 µl of 15% CHAPS was added to the protein fractions immediately. The protein was dialyzed against 10 mM Tris pH=7.5, 500 mM NaCl and 1.1% CHAPS and used for crystallization experiments immediately.

d. Purification of the NifA-like protein from Aquifex aeolicus

Purification of the NifA-like protein was done as described by Studholme *et al.*: Cells were broken up by lysozyme treatment and sonication for 3x1 m. Complete© protease inhibitor (Roche) was given to the cell extract and it was centrifuged at 15,000 g for 20 m. The supernatant was heated to 80 °C for 10 m and the suspension was centrifuged at 15,000 g for 20 m. The supernatant was loaded on a heparin column and eluted with the linear gradient of 10 mM NaH_2PO_4 pH=7.0, 2 M NaCl and 5% glycerol. Fractions containing the NifA-like protein were loaded on a Ni-NTA column and eluted with 50 mM NaH_2PO_4 pH=7.0, 500 mM NaCl, 5% glycerol and 1 M imidazole. The buffer was exchanged to 10 mM H/Na acetate pH=4.5 and 50 mM NaCl.

e. Purification and thrombin cleavage of ZraR-CC

The cells were lysed by the addition of lysozyme followed by sonication on ice for 3x1 m. The extract was centrifuged at 20,000 g for 20 m and the supernatant was loaded on a Ni-affinity column and eluted with the linear gradient of degassed 50 mM NaH_2PO_4 pH=8.0, 500 mM NaCl, 5% glycerol and 1 M imidazole. The sample was dialyzed against 500 mM NaCl, 10 mM Tris-HCl pH=8.0, 2 mM DTT and 2 mM $CaCl_2$ and digested with 10 U/mg bovine thrombin overnight at room temperature to remove the His-tag. After treatment with thrombin the sample was dialyzed against 20 mM CHES (2-(cyclohexylamino) ethane sulfonic acid) pH=9.5, 50 mM NaCl, 2 mM DTT, 0.1 mM EDTA and applied to a heparin column. The protein was eluted with the gradient of 20 mM CHES pH=9.5, 2 M NaCl, 2 mM DTT, 0.1 mM EDTA. Finally, the buffer of the protein sample was exchanged to 50 mM NaCl, 10 mM CHES pH=9.5, 2 mM DTT and 0.1 mM EDTA, which was supplemented with 20% glycerol and the protein was aliquoted and frozen in liquid nitrogen when crystallization experiments were not performed immediately.

5. Dynamic light scattering experiments

For dynamic light scattering (DLS) experiments the protein solution was diluted to ~0.5-1.0 mg/ml. The protein was applied to the cuvette without filtering and ~20-25 measurements were taken with a DynaPro dynamic light scattering machine for each protein solution.

6. Aluminium-trapping

The Al-trapping experiments were done as described by Chaney *et al.*: The buffer of 5 mM $MgCl_2$, 0.4 mM ADP, 10 mM NaF, 0.4 mM $AlCl_3$ and 10 mM Tris pH=8.0 was added to the solution of 0.7 mg/ml PspF-ΔHTH C146S and the mixture was incubated at 30 °C for 10 m.

7. Crystallization

For initial crystallization experiments, the sparse matrix screens Wizard I, Wizard II, Cryo I and Cryo II from Emerald BioStructures, the Crystal Screen and Crystal Screen 2 from Hampton Research were tried. The protein solutions were crystallized using sitting-drop vapour-diffusion. For the initial screens, 1 µl of protein solution was mixed with 1 µl of reservoir solution.

a. Crystallization of the PspF constructs

Crystallization trials on the PspF-ΔHTH with and without the His-tag, the PspF-ΔHTH C146S, the PspF-C and the PspF-C C146S constructs were set up. Additionally, crystallization of the PspF-ΔHTH was attempted in the Al-trapped form, with ADP and AMP-PNP. The protein was concentrated to 5-15 mg/ml and crystallization trays were set up at room temperature (20 °C). Crystals of the PspF-ΔHTH construct with the His-tag grew in 100 mM imidazole pH=7.5 and 800 mM Na-acetate. The PspF-ΔHTH construct without the His-tag gave crystals in 20% PEG-1K (polyethylene glycol), 100 mM Na/K phosphate pH=6.2, 200 mM NaCl and in 40% PEG-300, 100 mM cacodylate pH=6.5 and 200 mM Ca-acetate. The Additive, Detergent and PEG/Ion Screens from Hampton Research and two-dimensional screening techniques were tried in order to refine the crystallization conditions. The final buffer was 13% PEG-3350, 100 mM Na/K phosphate pH= 6.0 and 200 mM NaCl.

b. Crystallization of PspA

The protein was concentrated to 5.0 mg/ml in a buffer containing 10 mM Tris pH=7.5, 500 M NaCl and 1.1% CHAPS. In some experiments the CHAPS was replaced with β-octyl-glycoside (β-OG) . Crystallization trays were set up at 4 °C.

c. Crystallization of the NifA-like protein

The protein was concentrated to 10 mg/ml in 10 mM H/Na Acetate pH=4.5, 50 mM NaCl and crystallization trays were set up at 18 °C.

d. Crystallization of ZraR

The ZraR protein was concentrated to 10-12 mg/ml in 50 mM NaCl, 10 mM CHES pH=9.5, 2 mM DTT and 0.1 mM EDTA. Initially, crystals were obtained at 18 °C from 15% ethanol, 0.1 M Hepes pH=7.2, 200 mM $MgCl_2$ being a condition of the Wizard II sparse matrix screen. The conditions could not be optimized until it was realised that isopropanol, probably an additive in the ethanol used, was essential for crystallization. Subsequently crystals were obtained at 18 °C in 100 mM Hepes pH=7.5, 10% isopropanol, and 200 mM NaCl. With the Detergent and Additive screens from Hampton Research the crystallization conditions were further optimized with the best crystals obtained using a well solution containing 100 mM Hepes pH=8.2, 9% isopropanol, 3% methanol and 200 mM NaCl. The protein/precipitant ratio was optimized to 1:0.6. The Se-Met containing protein was crystallized in a N_2 atmosphere using a glove-bag.

8. Heavy-metal soaking experiments

For heavy metal soaking experiments of the ZraR crystals, 15 solutions of the Hg Screen from Hampton Research and $KAu(CN)_2$, K_2PtCl_4, $Pt(NH_3)_2Cl_2$ and $SmCl_2$ solutions were tried. The native crystals were transferred to the crystallization buffer containing ~1 mM of the heavy metal ion for 5, 10, 30 m and overnight. The optimized solution contained 1 mM ethyl-mercuric phosphate (EMP) in 100 mM Hepes pH=8.2, 9% isopropanol, 3% methanol and 200 mM NaCl with 5 m soaking time. The crystals were back-soaked in the cryo buffer (25% isopropanol, 20% PEG-200, 2% glycerol and 100 mM Hepes pH=8.2) for 5 m. The crystals were then either flash-cooled in nitrogen stream at 100 K, or plunged into, and stored in, liquid nitrogen.

9. Data collection and processing

Data from the native and Se-Met ZraR crystals at the peak and inflection point wavelengths were collected at the ID-14-4 beamline at the ESRF in Grenoble using an ADSC Quantum-4 detector. The wavelengths were chosen on the basis of the X-ray absorption fluorescence spectrum. Data on the mercury derivative ZraR crystals were collected at the X13 beamline at the EMBL in Hamburg using MAR165 CCD detector. The data were processed and reduced using the programs DENZO, SCALEPACK (Otwinowski and Minor, 1997) and TRUNCATE from the CCP4 program suite (Collaborative Computational Project, Number 4, 1994).

10. Phasing

SHELXD (Schneider and Sheldrick, 2002), SnB (Weeks and Miller, 1999), CRUNCH (de Graaf *et al.*, 2001) and SOLVE (Terwilliger, 2002) were used to search for the Hg and Se sites using SAS, SIRAS or MAD strategies. The self-rotation function was calculated with POLARRFN (Collaborative Computational Project, Number 4, 1994). The Hg positions were refined using SOLVE, the initial map was solvent flattened and NCS averaged with RESOLVE. SHARP (La Fortelle and Bricogne, 1997) was used to find additional Se sites from SAS data at the peak wavelength with the initial phases from the mercury derivative.

11. Model building and refinement

The model building was conducted with the program O (Jones *et al.*, 1991). Before refinement, 5% of all data were set aside for cross-validation. The initial model was refined with CNS (Brünger *et al.*, 1998) to an overall R-factor of 34% (R_{free}=37%), which was followed by rebuilding and TLS refinement cycles with Refmac (Murshudow *et al.*, 1997). Automated water building was carried out into the $|F_o|-|F_c|$ map with the CNS package. After adding the water molecules one final refinement cycle was done with Refmac to R_{work}=25.2% and R_{free}=30.8%. Structural homologues were identified with DALI (Holm and Sander, 1993). The Ramachandran plot was calculated with PROCHECK (Laskowski *et al.*, 1993). The figures were prepared with Molscript (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997)

Results

1. Overexpression, purification and crystallization of the PspF constructs

The PspF-ΔHTH protein was overproduced as detailed under Materials and Methods. From 1 L of LB media about 20 mg of purified protein was recovered. DLS experiments were carried out to optimize the pH and NaCl concentration of the protein buffer prior to crystallization experiments (Fig. 5). Crystallization trials on the construct with the His-tag gave small, non-diffracting crystals in 100 mM imidazole pH=7.5 and 800 mM Na-acetate. Extensive screening around this condition did not improve crystal quality. AMP-PNP and ADP with Mg^{2+} were tried to obtain crystals of better quality without any success.

One of the reasons of the poor crystal quality may be the presence of the truncated HTH domain at the C-terminus of the protein. To obtain a more compact protein, the PspF-C construct was designed and produced for crystallization experiments. Crystals of this construct were not obtained after extensive screenings.

Incubation of PspF with the transition state ATP analogue ADP-AlF_x ("Al-trapping") allows the activator to bind the σ^{54} factor, which is an indication of the existence of a transient conformation (Chaney *et al.*, 2001). Crystallization trays with the "trapped" PspF-ΔHTH protein were set up as well, but crystals were not obtained.

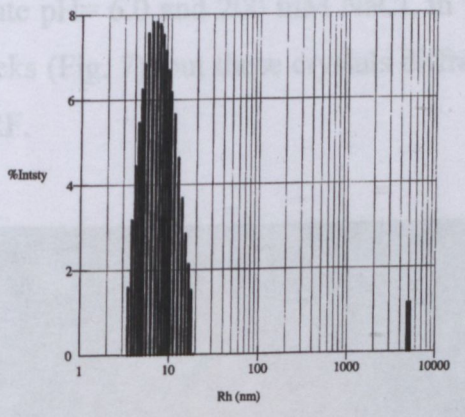


Fig. 5. DLS histogram of the PspF-ΔHTH protein in 10 mM Tris pH=7.5 and 50 mM NaCl. The first peak at 7.9 nm corresponds to a multimeric complex consisting of about 12 monomers, the second is probably dust contamination.

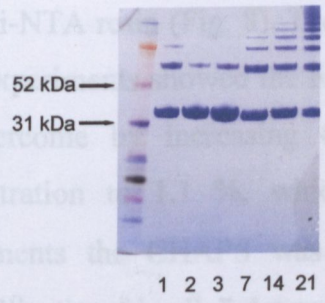


Fig. 6. SDS-PAGE gel on PspF-ΔHTH that shows covalently cross-linked multimers. The protein was incubated at room temperature for three weeks and samples were taken on a regular basis. The incubation time is given in days.

Incubation of the protein solution at room temperature proved that the construct has a tendency to make covalent cross-links between the monomers (Fig. 6), which may happen in the crystallization drop as well and inhibit the crystallization or influence the crystal quality. The formation of covalently cross-linked oligomers may be due to two possible reasons; either the three cysteins present in the sequence oxidize or the His-tag at the N-terminus makes cross-links between the monomers.

To overcome this problem, the C146S mutation was introduced into the PspF-C and PspF-ΔHTH constructs and crystallization trays were set up, but the mutated proteins did not give crystals even under the known conditions mentioned above. To address the second problem, the His-tag of the PspF-ΔHTH construct was removed by thrombin treatment. Small, sphere-like crystals of this construct grew in 20% PEG-1K, 100 mM Na/K phosphate pH=6.2, 200 mM NaCl and in 40% PEG-300, 100 mM cacodylate pH=6.5 and 200 mM Ca-acetate. The conditions were optimized to 13% PEG-3350, 100 mM Na/K phosphate pH= 6.0 and 200 mM NaCl, in which big, cubic crystals grew over a period of two weeks (Fig. 7), but these crystals diffracted only to ~10 Å on the ID-14-4 beamline at the ESRF.

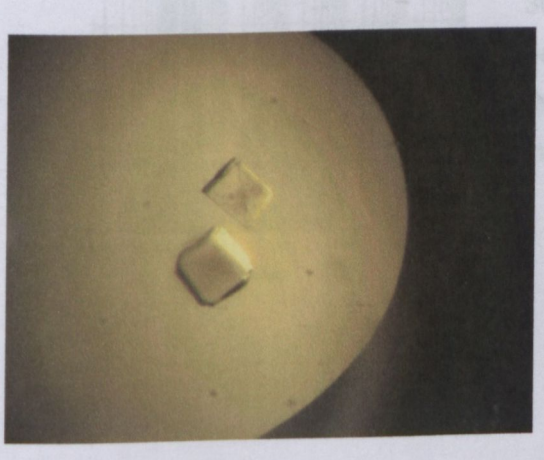


Fig. 7. Native crystals of the PspF-ΔHTH protein. The crystal size is aprroximately 200 x 200 x 100 μm.

2. Overexpression, purification and crystallization of PspA

The PspA protein was successfully overproduced in *E. coli B834 (DE3)* cells and purified on a Ni-NTA resin (Fig. 8). The amount of the purified protein reached 3-4 mg/L culture. DLS experiments showed the high aggregation tendency of the protein (Fig. 9) that could be overcome by increasing the NaCl concentration to 500 mM and the CHAPS concentration to 1.1 %, which buffer was used for crystallization trials. In some experiments the CHAPS was replaced with β-OG, which is known to be a more "crystallization friendly" detergent, however, the PspA protein was partially aggregated in the lack of CHAPS. Crystals were not obtained in either of the trays.

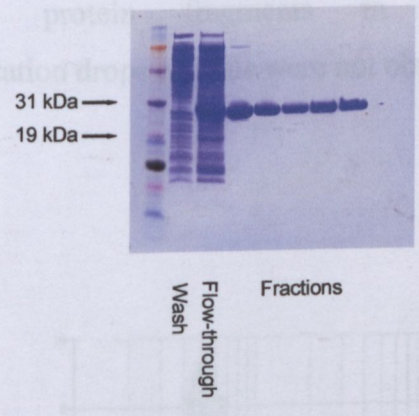


Fig. 8. SDS-PAGE gel on the PspA protein. Samples were taken from different fractions coming from the Ni-affinity purification step.

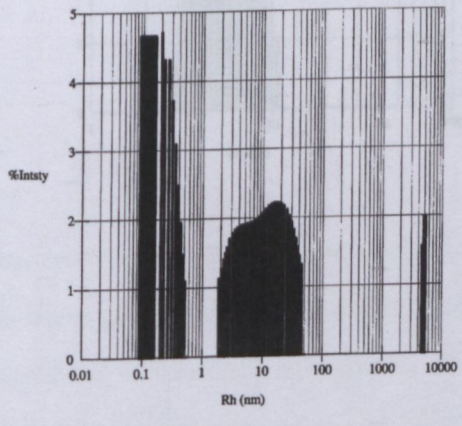


Fig. 9. DLS histogram of the PspA protein. The buffer contains 10 mM Tris pH=7.5, 300 mM NaCl and 0.01 % CHAPS. The first peak corresponds to CHAPS, the second represents the aggregated protein particles, the third is dust contamination.

3. Overexpression, purification and crystallization of the NifA-like protein

Since the gene of the NifA-like protein contains some amino acid codes used rarely by *E. coli*, the protein could be successfully overexpressed in *E. coli BL21 Codon Plus* cells only. After the purification, about 5 mg protein was recovered. Despite the use of powerful protease inhibitors during the purification the protein was degraded (Fig. 10). The protein buffer was optimized to 10 mM H/Na acetate pH=4.5 and 50 mM NaCl with DLS prior to crystallization (Fig. 11). Probably due to the presence of different protein fragments in the crystallization drops crystals were not obtained.

Fig. 10. SDS-PAGE gel on the NifA-like protein after the Ni- and Heparin-affinity purification steps. The gel shows many degradation products that could not be removed during the purification.

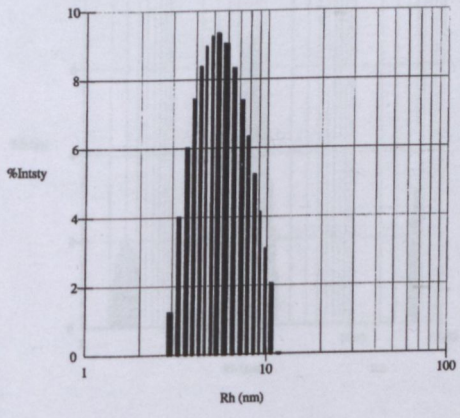


Fig. 11. DLS histogram of the NifA-like protein, which was solved in H/Na acetate pH=4.5 and 50 mM NaCl. The protein shows a monomodal distribution with a polydispersity index of 28.2%. The median is at 4.47 nm, which corresponds to a molecular weight of 109 kDa indicating a dimer in the solution.

4. Overexpression, purification, crystallization and phasing of ZraR-CC

ZraR-CC was overexpressed and purified as described above with a yield of about 5 mg/L culture pure protein (Fig. 12). The protein buffer was optimized to 10 mM CHES pH=9.5, 50 mM NaCl, 1 mM DTT and 0.1 mM EDTA in which DLS measurements indicated the presence of minimal amount of aggregated protein (Fig. 13).

Crystals grew overnight to a maximum length of ~300 μm in 100 mM Hepes pH=8.2, 9% isopropanol, 3% methanol and 200 mM NaCl; the average approximate dimensions were 150 x 80 x 50 μm (Fig. 14). The crystals belong to space group $p222_1$. The volume of the unit cell is 2.31 x 10^6 $Å^3$, which is consistent with a hexamer in the asymmetric unit (V_m =2.91 $Å^3$ Da^{-1}) as confirmed by the self-rotation function. The best native crystal diffracted to 3.0 Å (Fig. 15).

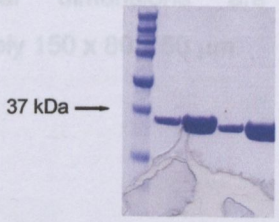


Fig. 12. SDS-PAGE gel showing the ZraR protein at two concentrations (0.1 and 0.5 mg/ml) after purification before (1) and after (2) thrombin cleavage.

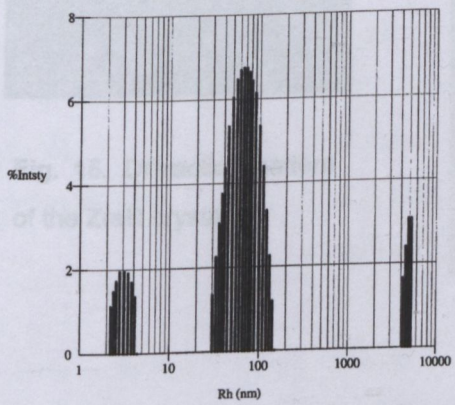


Fig. 13. DLS histogram of ZraR-CC ready for crystallization experiments. The protein is in 10 mM CHES pH=9.5 and 50 mM NaCl. The first peak (2.8 nm, 31.4 kDa) with a poydispersity index of 17.4% represents the monomer. Some of the protein is aggregated (less than 0.1% in mass) which is indicated by the second peak.

The Se-Met crystal diffracted initially (at the peak wavelength) to 3.2 Å but, due to radiation damage, the inflection point data were processed only to 3.4 Å. Probably as a consequence of the radiation damage, SAS and two-wavelength MAD strategies using different programs failed to give solutions for the selenium sites. The Se-Met crystal was non-isomorphous with the native crystal (Table 1.) and consequently the isomorphous differences between native and Se-Met data sets also failed to yield the selenium positions.

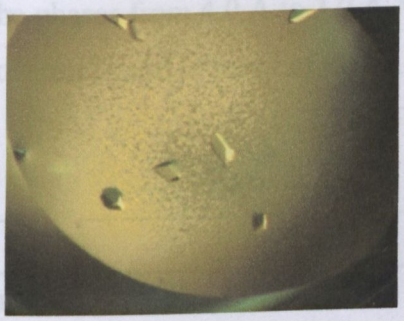
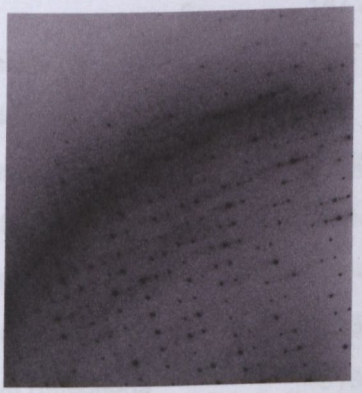


Fig. 14. Native crystals of ZraR central and DNA-binding domains. The typical dimensions are approximately 150 x 80 x 50 μm.



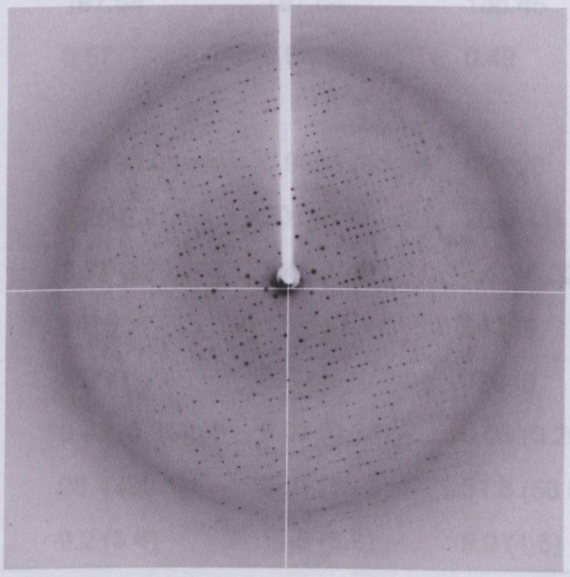


Fig. 15. Diffraction pattern of the ZraR crystals.

Soaking the native crystals in 1 mM ethyl-mercuric phosphate for 5 m gave Hg derivative crystals diffracting to 4.0 Å. All the six initial heavy atom positions were found from the SAS data with SHELXD, the sites were refined with SOLVE. Data from the native crystal were used to calculate SIRAS phases with an overall figure of merit of 0.17. With help of the initial phases from the Hg SIRAS phasing all the 36 Se sites were located using an SAS strategy with SOLVE and SHARP. The maps produced from SOLVE and SHARP were noisy and did not reveal any recognizable protein features. Surprisingly,

phase combination from the two different sources did not improve the map quality, so only the SIRAS phases from the Hg sites were used to calculate an electron density map. Solvent flattening and sixfold NCS operators determined from the Se sites were applied to extend the phases from 4.0 Å to 3.0 Å, which yielded a map with recognisable features and with an overall figure of merit of 0.45.

	Native	Se-Met crystal		Hg derivative
Crystal data				
Space group		P2221		
Unit cell parameters (Å)				
a	107.44	109.43	109.72	106.88
b	114.74	112.00	111.97	113.94
c	187.26	187.84	187.84	186.46
Mosaicity (°)	0.73	0.57	0.68	0.49
Data collection				
Wavelength (Å)	0.9806	0.9795 (peak)	0.9806 (edge)	0.8020
Resolution range (Å)	30.0-3.0	30.0-3.2	30.0-3.4	20.0-4.0
Highest resolution shell (Å)	3.11-3.00	3.25-3.20	3.45-3.40	4.10-4.00
Observed reflections	173549	142605	118851	86339
Unique reflections	46186	38771	32788	17045
R_{merge}† (%)	0.06 (0.44)	0.08 (0.35)	0.07 (0.35)	0.08 (0.26)
Completeness (%)	97.5 (87.8)	98.4 (90.1)	98.1 (88.5)	81.8 (58.5)
I/σ	10.9 (2.6)	9.2 (3.4)	11.6 (3.3)	9.0 (4.5)

Table 1. Data collection statistics. Values in parenthesis are for the highest resolution shell. † $R_{merge} = \sum|I_i-(I_i)|/\sum(I_i)$, where I_i is the observed intensity and (I_i) is the average intensity over symmetry-equivalent intensities.

5. Crystal structure of ZraR

a. Overall structure of ZraR

The structure was refined to an overall R-factor of 25.2% (R_{free}= 30.8%). The crystallographic refinement statistics are summarized in Table 2. The structure reveals a hexameric architecture (Fig. 16A), which is typical for the AAA+ domain containing proteins. There are two hexamers each with crystallographic twofold symmetry about b (Fig. 17). The hexamers pack head-to-head with contacts on one side between the HTH domains and the other between the central domains. The DNA-binding HTH motifs do not have the non-crystallographic symmetry of the ATPase domains and form a half ring comprised of domains belonging to adjacent hexamers (Fig. 16B). The N-terminus, which would connect to the missing receiver domain, is on the opposite side of the hexameric ring from the HTH domains.

R values and B-factor		Model Geometry	
Resolution (Å)	20.0-3.0	R. m. s deviations	
Number of reflections		Bonds (Å)	0.020
For refinement	43710	Angles (°)	2.34
For test set	2290	B-factor (Å^2)	19.9
R_{cryst}† (%)	25.2 (32.2)	Ramachandran plot (%)	
R_{free}† (%)	30.8 (37.0)	Most favoured	76.5
Average B (Å^2)	58.4	Additionally allowed	21.3
No. of atoms in refinement		Generously allowed	2.2
Total	12594		
Waters added	102		

Table 2. Refinement statistics. Values in parenthesis are for the highest resolution shell. † R_{cryst} and $R_{free} = \sum|F_{obs}-(F_{calc})|/\sum(F_{obs})$, for reflections in the working and test sets, respectively.

5. Crystal structure of ZraR

a. Overall structure of ZraR

The structure was refined to an overall R-factor of 25.2% (R_{free}= 30.8%). The crystallographic refinement statistics are summarized in Table 2. The structure reveals a hexameric architecture (Fig. 16A), which is typical for the AAA+ domain containing proteins. There are two hexamers each with crystallographic twofold symmetry about b (Fig. 17). The hexamers pack head-to-head with contacts on one side between the HTH domains and the other between the central domains. The DNA-binding HTH motifs do not have the non-crystallographic symmetry of the ATPase domains and form a half ring comprised of domains belonging to adjacent hexamers (Fig. 16B). The N-terminus, which would connect to the missing receiver domain, is on the opposite side of the hexameric ring from the HTH domains.

R values and B-factor		Model Geometry	
Resolution (Å)	20.0-3.0	R. m. s deviations	
Number of reflections		Bonds (Å)	0.020
For refinement	43710	Angles (°)	2.34
For test set	2290	B-factor (Å^2)	19.9
R_{cryst}† (%)	25.2 (32.2)	Ramachandran plot (%)	
R_{free}† (%)	30.8 (37.0)	Most favoured	76.5
Average B (Å^2)	58.4	Additionally allowed	21.3
No. of atoms in refinement		Generously allowed	2.2
Total	12594		
Waters added	102		

Table 2. Refinement statistics. Values in parenthesis are for the highest resolution shell. † R_{cryst} and $R_{free} = \sum|F_{obs}-(F_{calc})|/\sum(F_{obs})$, for reflections in the working and test sets, respectively.

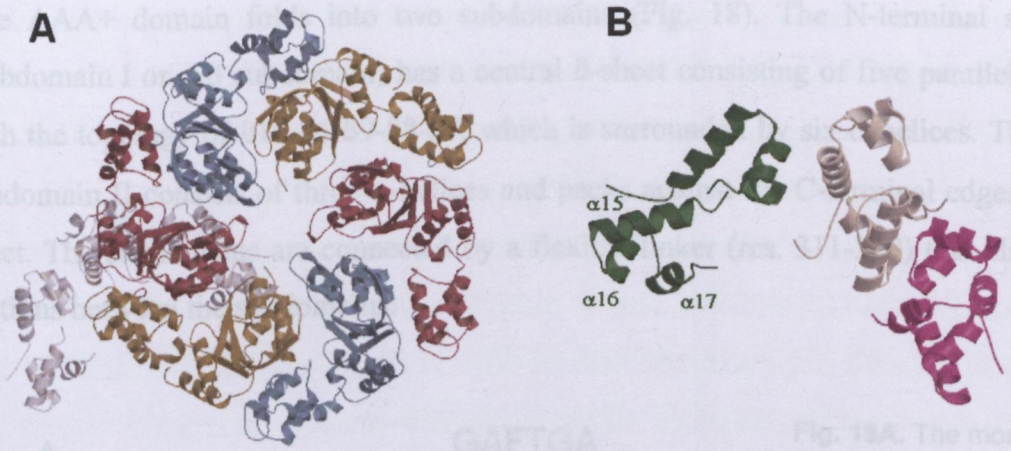


Fig. 16. Overall structure of the AAA and DNA-binding domains. (**A**) Three molecules (A in yellow, B in brown and C in cyan) of the asymmetric unit form a hexamer with crystallographic twofold symmetry about b. The remaining three molecules form a similar hexamer about a second crystallographic twofold. The position of the three dimers formed by the HTH domains underneath the hexamer is indicated in white. (**B**) Association of three dimers of HTH domains. Each domain belongs to different hexamer. Three of the six HTH domains are not visible in the electron density map.

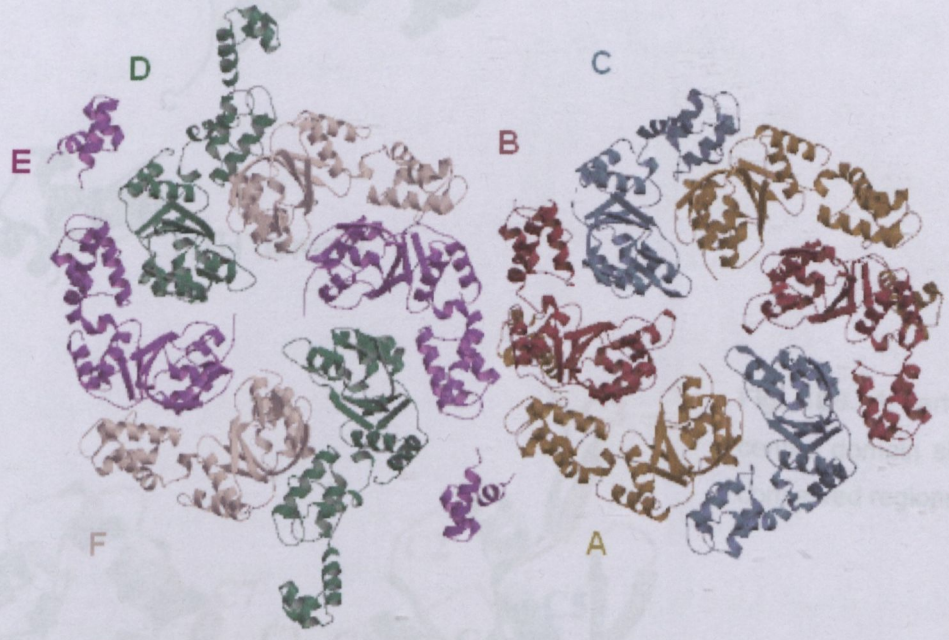


Fig. 17. Two hexamers of which six molecules labelled with different colours represent the crystallographic asymmetric unit. The DNA-binding domains of molecules A, D and E are visible in the electron density map.

b. Structure of the AAA domain

The AAA+ domain folds into two subdomains (Fig. 18). The N-terminal subdomain (subdomain I or α/β subdomain) has a central ß-sheet consisting of five parallel ß-strands with the topology of ß10-ß6-ß9-ß8-ß7, which is surrounded by six α-helices. The smaller subdomain II consists of three α-helices and packs against the C-terminal edges of the ß-sheet. The subdomains are connected by a flexible linker (res. 311-319) that allows large motions between the subdomains.

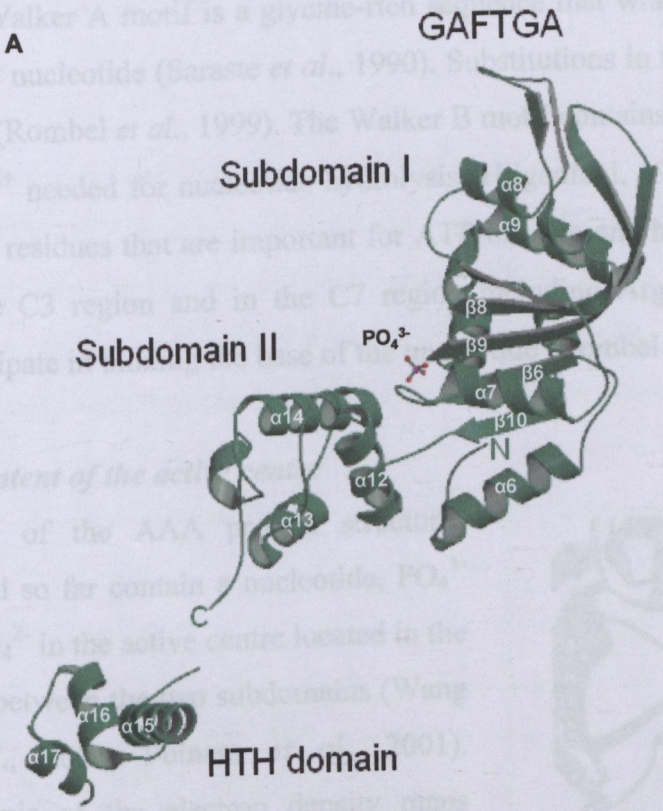


Fig. 18A. The most complete protomer in the asymmetric unit is molecule D. The HTH domain and the main chain of the loop containing the GAFTGA motif are visible as well. In the active centre of the molecule a PO_4^{3-} has been placed.

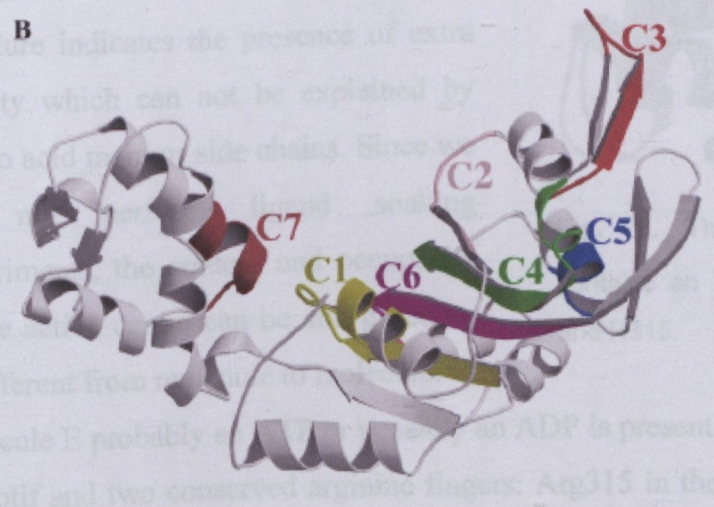


Fig. 18B. A cartoon of the central domain showing the conserved regions C1-C7.

The central domain is highly conserved in all known σ^{54}-dependent activators and can be subdivided into seven well-conserved regions (Fig. 4; Osuna *et al.*, 1997). The conserved D/ESELFGHE (residues 207-214) and GAFTGA (res. 216-221) motifs (Fig. 18, conserved region C3) that play an important role in the binding to the σ^{54} transcription factor (North *et al.*, 1996; Yan and Kustu, 1999; Bordes *et al.*, 2003; Rappas *et al.*, 2005), are part of α8 and a very flexible loop, respectively, and point towards the central hole of the hexamer. The domain contains the Walker A (res. 169-176, C1) and Walker B motifs (res. 235-242, C4) that participate in ATP binding and hydrolysis (Walker *et al.*, 1982). The Walker A motif is a glycine-rich sequence that wraps around the β-phosphate of the bound nucleotide (Saraste *et al.*, 1990). Substitutions in this motif abolish ATP binding in NtrC (Rombel *et al.*, 1999). The Walker B motif contains a conserved Asp that coordinates a Mg^{2+} needed for nucleotide hydrolysis (Hilgenfeld, 1995) and lies between ß8 and α9. Other residues that are important for ATP binding and hydrolysis include certain residues in the C3 region and in the C7 region including Arg359 (NtrC Arg358) of α14 that participate in binding the base of the nucleotide (Rombel *et al.*, 1999).

c. Content of the active centre

Many of the AAA protein structures solved so far contain a nucleotide, PO_4^{3-} or SO_4^{2-} in the active centre located in the cleft between the two subdomains (Wang *et al.*, 2001; Putnam *et al.*, 2001). Analysis of the electron density maps around the Walker A motif of the ZraR structure indicates the presence of extra density which can not be explained by amino acid main or side chains. Since we did not perform ligand soaking experiments, the content and occupancy of the active centre can be and probably is different from molecule to molecule. In molecule E probably an ATP or possibly an ADP is present, which is bound by the Walker A motif and two conserved arginine fingers: Arg315 in the interdomain loop and Arg359

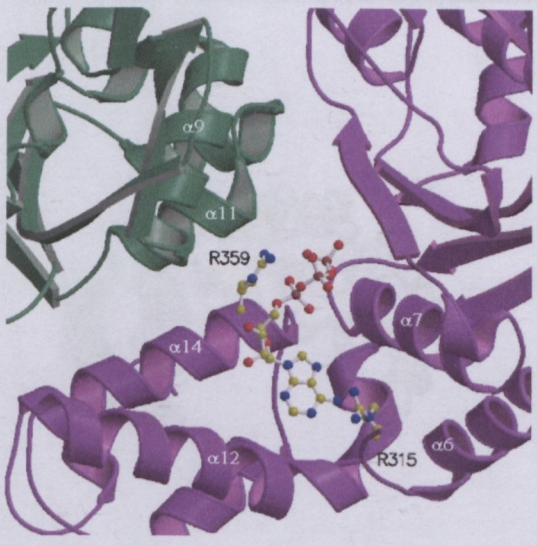


Fig. 19. The active center in molecule E contains an ATP molecule bound by R359 and R315.

in C7 (Fig. 19); whereas in the other five molecules only the presence of the γ-phosphate can be confirmed. The superposition of the six molecules using the N-terminal subdomain shows the conformational change (primarily a rotation) of the C-terminal subdomain depending on the content and occupancy of the active centre (Fig. 20). Molecule E is in the most open, whereas molecule C is in the most closed conformation. This observation contrasts with that for the HslU heat shock locus protein U (Wang *et al.*, 2001), the RuvB branch migration protein (Putnam *et al.*, 2001), and the p97 ATPase (Rouiller *et al.*, 2002), where bound ATP gives rise to a closed conformation. The difference in the content and occupancy suggests the presence of different nucleotides and conformational states in the crystal.

d. Structure of the DNA-binding domain

The HTH DNA-binding domain has three α-helices and is connected to the central domain by a flexible linker consisting of about fifteen amino acids, which allows independent motion relative to the AAA domain (Fig 16B, Fig. 18). The longest helix in the domain, α15, packs with that of another subunit of an adjacent hexamer with a twofold symmetry.

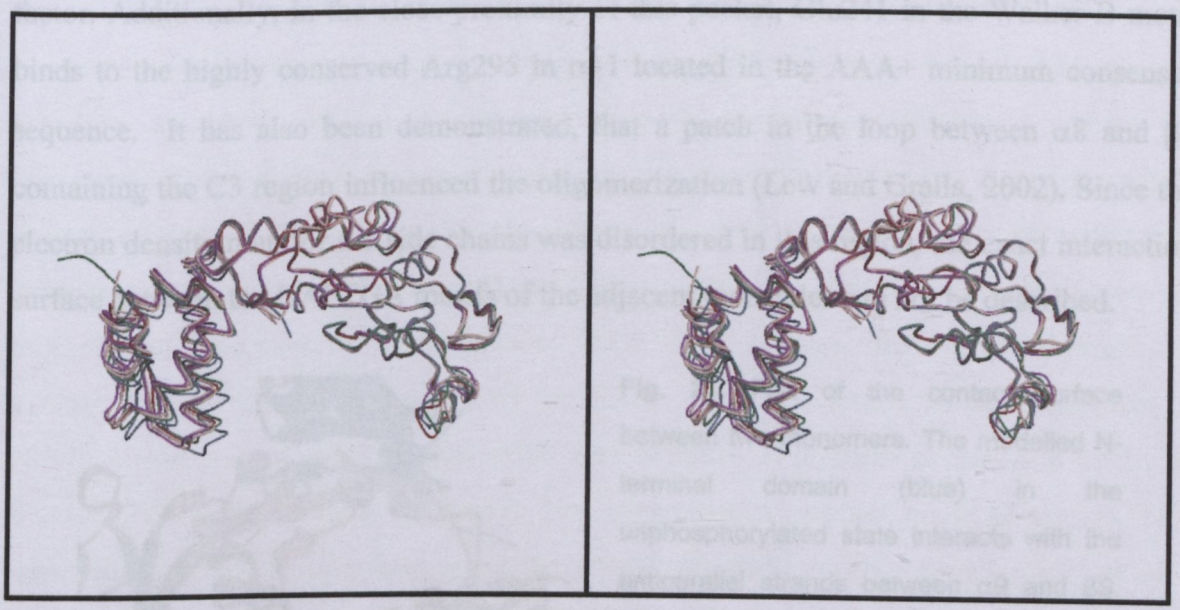


Fig. 20. A stereographic drawing of the superposition of Cα traces for the central domain of the six molecules. Molecule E (lila) is in the most open, whereas molecule C (cyan) is in the most closed conformation.

Helices α16-17 adopt a classic HTH DNA-binding motif. In the electron density map only three of the six HTH domains are visible that form a half ring, which is independent from the hexamer of the AAA domains due to the long linker between them. The HTH domains in this half ring belong to different hexameric rings and form dimers. These interactions stabilize the packing between independent hexameric rings. Though the HTH domains of NtrC bind to the DNA as dimers (Klose *et al.*, 1994), the structure of this dimer is not that observed in the crystal structure reported here. We suggest that the DNA-binding domains of neighbouring protomers in the same hexamer dimerize to bind to the DNA.

e. Contact surface between the monomers

The contact surface between the adjacent AAA+ domains in the hexameric ring is made up of both hydrophobic and hydrophilic interactions (Fig. 21). The C-terminal end of α14 forms a hydrophobic pocket with the C-terminus of α6, the following loop and the N-terminus of ß10 of the adjacent subunit. Salt bridges between α14 and α11 contribute to this interaction surface. The conserved Ala199, Ala200 and Leu201 in the turn between ß7 and α8 interacts with Leu248 and Val251 in α9 on the other side, which seems to be very important since this hydrophobic pocket is very close to the C3 region binding to the σ^{54} factor. Additionally, in the close proximity of this pocket, Glu241 in the Walker B motif binds to the highly conserved Arg295 in α11 located in the AAA+ minimum consensus sequence. It has also been demonstrated, that a patch in the loop between α8 and β8 containing the C3 region influenced the oligomerization (Lew and Gralla, 2002). Since the electron density map for the side chains was disordered in this region, the exact interaction surface between the GAFTGA motifs of the adjacent molecules can not be described.

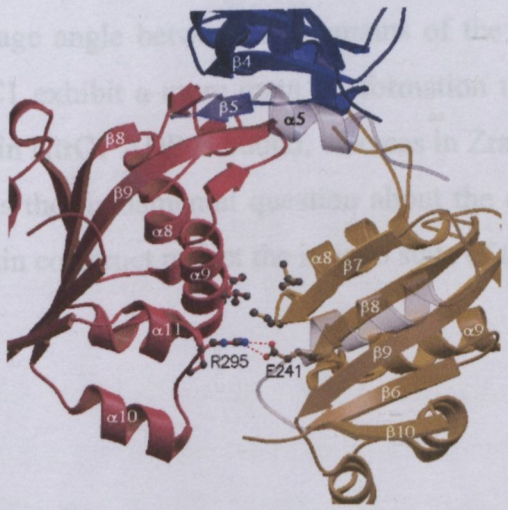


Fig. 21. Part of the contact surface between two monomers. The modelled N-terminal domain (blue) in the unphosphorylated state interacts with the antiparallel strands between α9 and ß9. The release of this contact probably allows the rotation of α9 and the formation of a hydrophobic pocket between the two molecules. The formation of the Arg295-Glu241 salt bridge is probably the trigger for the ATP cleavage.

Discussion

1. Oligomerisation of enhancer binding proteins

Although the structures of isolated regulatory and DNA-binding domains have been solved (Kern *et al.*, 1999; Meyer *et al.*, 2001; Park *et al.*, 2002), there was no structural data available on the oligomerisation and ATPase domain of EBPs. The crystal structure of NtrC1 from *Aquifex aeolicus* (Lee *et al.*, 2003) and the cryogenic electron microscopy structure at 20 Å of PspF in complex with σ^{54} from *Escherichia Coli* (Rappas *et al.*, 2005) in parallel with our work gave us insights into the oligomerisation and activation of EBPs.

EBPs are usually dimeric in their inactive state (Klose *et al.*, 1994; Rippe *et* al., 1998; Schumacher *et al.*, 2004) and need to form higher oligomers for ATPase activity (Wyman *et al.*, 1997; Rappas *et al.*, 2005). EBPs bind to upstream activator sequences (UAS) that are located distant from the transcriptional start site. As described in Chapter 1, EBP-σ^{54} interactions require DNA loop formation between the UAS and promoter sites (Fig. 1). UAS binding was shown to promote higher oligomer formation of some EBPs and to increase ATP hydrolysis rate (Jovanovic *et al.*, 1997; Wyman *et al.*, 1997).

The AAA+ domain of NtrC1, which is active in the absence of the receiver domain, crystallized as a ring-shaped heptamer (Fig. 22). The inactive construct of the protein containing the receiver and AAA+ domains formed dimers in the crystal. The dimeric structure suggested that the N-terminal regulatory domain controls NtrC1 activity by hindering higher oligomer formation: extensive contacts between the two domains prevent oligomerisation of the ATPase domain (Lee *at al.*, 2003). The reason for the difference between the oligomeric states of ZraR and NtrC1 is not clear, though the consequences can be explained from structural point of view: in the NtrC1 heptamer the ring is flatter and the average angle between subdomains of the AAA+ domain is larger, so the monomers of NtrC1 exhibit a more open conformation than in ZraR. A possible explanation might be that in NtrC1 ADP is bound, whereas in ZraR phosphate or sulphate might be present. This raises the fundamental question about the crystal structures: does a crystal structure of a certain construct reflect the in vivo state of the protein?

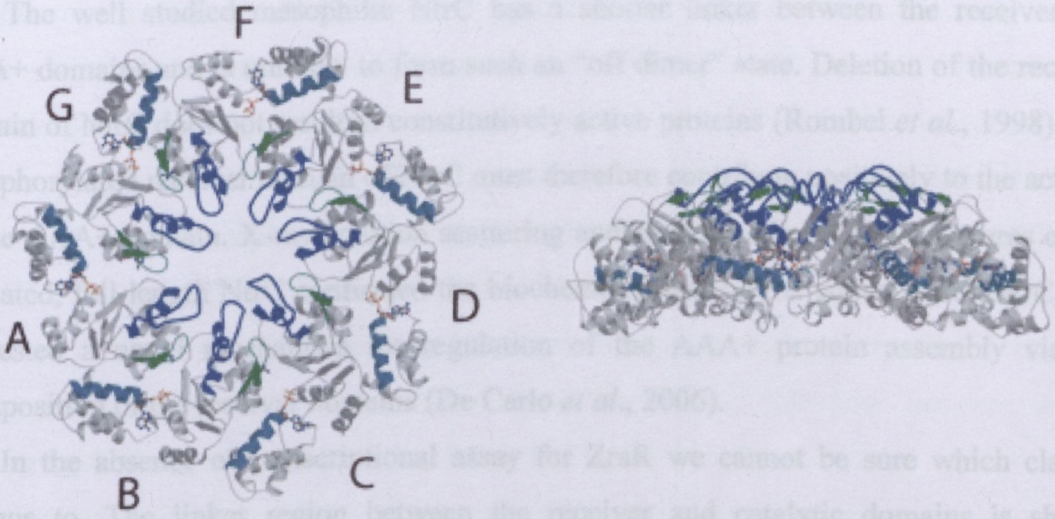


Fig. 22. Structure of the NtrC1 heptamer. At *left* is a top view illustrating how the protomers pack in the context of the heptamer. Each protomer is labeled with a letter, A through G. The GAFTGA insertion (blue) and β-hairpin (green) comprise the pore in the heptamer. At *right* is a side view (Lee *et al.*, 2003).

2. Domain structure and regulation of EBPs

Most EBPs consist of three domains, an N-terminal regulatory domain, the AAA+ domain that is often sufficient for transcription of σ^{54}-dependent promoters in vivo and in vitro, and the C-terminal HTH domain.

There are at least two classes of σ^{54}-dependent activators, depending on whether the N-terminal domain negatively or positively regulates the transcriptional activity (Park *et al.*, 2002). The receiver domains of NtrC1 from *Aquifex aeolicus* (Lee *et al.*, 2003) and DctD from Sinorhizobium meliloti (Park et al., 2002) regulate the activity of the AAA+ domain negatively as shown by the constitutive activity of those proteins when their receiver domain is deleted. The linker region between the DctD and NtrC1 receiver and AAA+ domains form a coiled-coil structure that maintain the proteins in an "off dimer" assembly (Doucleff *et al.*, 2005; Park *et al*, 2002). Phosphorylation of the N-terminal domains by their cognate sensor histidine kinases disrupts the coiled-coil structure by providing a different dimer interface, thereby shifting the relative orientation between N-terminal and AAA+ domain and allowing the formation of higher oligomers between

AAA+ domains. This type of regulation seems specific to a number of EBPs that show the sequences predicted to form coiled-coil structure between receiver and AAA+ domains.

The well studied mesophilic NtrC has a shorter linker between the receiver and AAA+ domains and is unlikely to form such an "off dimer" state. Deletion of the receiver domain of NtrC does not result in constitutively active proteins (Rombel *et al.*, 1998). The phosphorylated receiver domain of NtrC must therefore contribute positively to the activity of the AAA+ domain. X-ray solution scattering and electron microscopy structures of the activated, full-length NtrC confirmed the biochemical data on the activation of NtrC, and suggested a novel mechanism for regulation of the AAA+ protein assembly via the juxtaposition of the receiver domains (De Carlo *et al.*, 2006).

In the absence of transcriptional assay for ZraR we cannot be sure which class it belongs to. The linker region between the receiver and catalytic domains is shorter compared to DctD, which reduces the likelihood of a coiled-coil dimerisation determinant in the linker as observed for DctD. Like NtrC, ZraR has Ala at position 83 rather than Gly found in DctD and NtrC1. These factors suggest that ZraR more likely belongs to the NtrC subgroup rather than DctD and NtrC1.

The DNA binding domains for EBPs allow for specific UAS recognition. The HTH DNA binding domains of NtrC (Pelton *et al.*, 1999) and ZraR form dimers involving an alpha helix. Dimerisation does not occur in ZraR between DNA-binding domains from the same hexameric ring in the crystal structure. This could be because a side by side arrangement of protomers in the hexameric ring, coupled with the head to head arrangement of DNA-binding domains and constrained by a linker of only 15 residues makes this unfavourable. Thus, interhexameric interactions between two hexamers allow the existence of larger oligomers in solution, possibly dodecamers. Interestingly, dodecamers have been observed in cryoelectronmicrographs of PspF activated by ADP-AlF_x (Schumacher *et al*, 2004). So the functional interface between the AAA+ and DNA-binding domains is not evident, and it is not clear if such a dimerisation of DNA-binding domains is of physiological importance.

3. Structural elements of the AAA+ domain

The sequence similarities between AAA+ domains of different types of AAA+ members are low (Neuwald *et al.*, 1999). The α/β subdomains of the AAA+ domains are generally better conserved than the residues of the α-helical domain. The N-terminal α/β subdomain of AAA+ proteins is a typical Rossman Fold and conserved in EBPs. The highest sequence conservation between AAA+ members are in their Walker A and Walker B motifs (conserved region C1 and C4, respectively), the GAFTGA motif of σ^{54} activators (C3), and to a lesser extent in their sensor I (C6) and sensor II motifs (C7; Fig. 4 and 18B). The diversity in AAA+ domains comes from the number and position of the five or more α-helices that connect the strands of the central β-sheet and insertions between strands (Hanson and Whiteheart, 2005).

Compared to other AAA+ proteins the EBP α-helical subdomain reveals no major differences with respect to the α-helical arrangement. This subdomain is composed of several α-helices and lies above the wide end of the N-terminal wedge and forms a partial lid over the nucleotide binding site (Fig. 19).

The AAA+ domain of ZraR shows the highest structural similarity to PspF (Rappas *et al.*, 2005) and NtrC1 (Lee *et al.*, 2003). Pairwise overlays of the 232 α-carbon atoms that are structurally defined in all three structures have root mean square deviations of PspF-ZraR 4.3 Å, NtrC1-ZraR 1.5 Å, and NtrC1-PspF 4.2 Å. The differences mostly reflect the different orientations of the two subdomains. However, as illustrated on Fig. 20, the orientation of the α-helical subdomain of ZraR is slightly different in the six molecules of the asymmetric unit as well, with an rmsd of 0.3 Å.

4. Modelling of the unphosphorylated full-length protein

The projection structure of the full-length HupR protein, which is the regulator of the [NiFe] hydrogenase synthesis and belongs to the NtrC family, has been determined at 9 Å resolution. The N-terminal domain of NtrC was compared to the map and the two-dimensional orientation determined (Fig. 23A, Vénien-Bryan *et al.*, 2000). Taking this 2D comparison we have modelled the full-length ZraR with the receiver domain of NtrC (Fig. 23B).

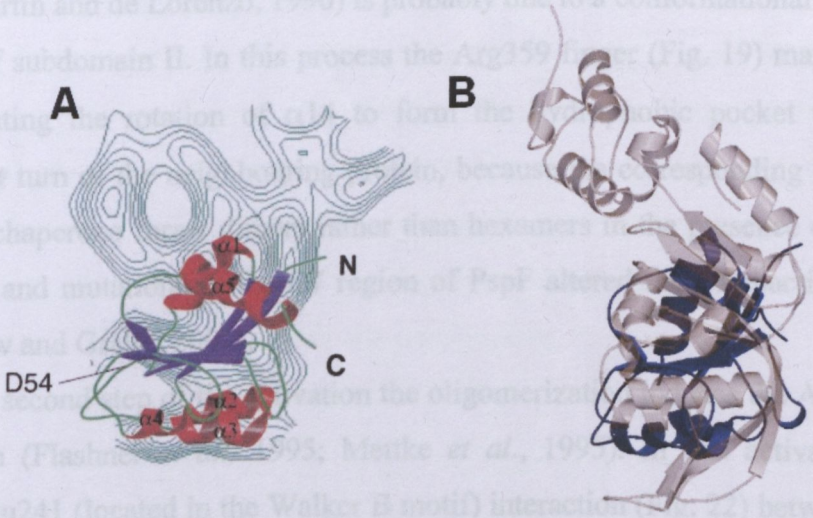


Fig. 23. **(A)** Projection structure of the full-length HupR protein at 9 Å resolution to which the N-terminal domain of NtrC was compared (Venien-Bryan *et al.*, 2000). (**B**) Crystal structure of ZraR to which the N-terminal domain of NtrC (blue) has been modelled based on the 9 Å projection map.

5. Mechanism of signal propagation

The unphosphorylated form of NtrC (and ZraR) is a dimer in solution and can bind both DNA and ATP (Rombel *et al.*, 1999). The major dimerization determinants are in the DNA-binding domain (Klose *et al.*, 1994), however, the AAA domain is responsible for the oligomerization of the protein when the N-terminal domain is phosphorylated (Flashner *et al.*, 1995). It has been shown, that phosphorylation at Asp54 of NtrC induces a large conformational change in the receiver domain, involving a displacement of β4-5 and α3-4 that exposes the hydrophobic surface of α4 (Leu87, Ala90, Val91) to transmit the signal to the AAA+ domain of the adjacent molecule in the dimer (Kern *et al.*, 1999; Nohaile *et al.*, 1997; Hwang *et al.*, 1999; Lee *et al.*, 2000). It has also been shown, that in the unphosphorylated state the receiver domain interacts with the AAA domain (Fiedler and Weiss, 1995). Based on our model of the full-length protein in the unphosphorylated state, when Asp54 is not phosphorylated, ß5 of the receiver domain joins the antiparallel ß-strands of the flexible loop between α9 and ß9 of the central ß-sheet (Fig. 21). This contact probably prevents the interaction surface on α9 described above being exposed. Upon phosphorylation, ß5 is replaced by α4; this contact is released, which we suggest induces the rotation of α9 allowing the hydrophobic surface to contact the neighbouring subunit

and hence forming the hexamer. The stimulation of oligomerization by ATP-binding (Perez-Martin and de Lorenzo, 1996) is probably due to a conformational change involving rotation of subdomain II. In this process the Arg359 finger (Fig. 19) may have a key role by facilitating the rotation of α14 to form the hydrophobic pocket with α6 and the subsequent turn of the neighbouring protein, because the corresponding R332A mutant in the ClpB chaperone forms dimers rather than hexamers in the presence of ATP (Mogk *et al.*, 2003) and mutations in the C7 region of PspF altered the compaction of the protein dimer (Lew and Gralla, 2002).

In the second step of the activation the oligomerization induces the ATPase activity of the protein (Flashner *et al.*, 1995; Mettke *et al.*, 1995). In this activation process the Arg295-Glu241 (located in the Walker B motif) interaction (Fig. 22) between two adjacent molecules plays a critical role, since the NtrC R295C mutant protein shows no ATPase activity (North *et al.*, 1996) and the corresponding R174 in RuvB plays an intermolecular catalytic role (Hishida et al, 2004). By analogy to the mechanism of HslU (Wang *et al.*, 2001) and RuvB (Putnam *et al.*, 2001), the cleavage of ATP and the PO_4^{3-} release shifts α7 towards the C-terminus by ~3Å and causes the back rotation of subdomain II, which disrupts the interaction with the neighbouring subunit and may result in transient dissociation of the hexamer.

6. Modelling of the full-length protein in the active state

The presence of the N-terminal domain in the wild type protein is essential for the ATPase activity but not for the ATP-binding of the NtrC family (Drummond *et al.*, 1990). It has been demonstrated, that the N-terminal domain of the N-ethylmaleimide-sensitive factor rotates and twists upon ATP hydrolysis (Yu *et al.*, 1999). Additionally, the N-terminal domain of p97 became ordered and visible in the cleft between the two subdomains only in the p97-ADP-AlF_x complex but not with AMP-PNP, ADP or in the empty state in the electron microscopic map (Rouiller *et al.*, 2002).

Studies with an iron chelate reagent conjugated to the receiver domain of NtrC suggest that the phosphorylation-dependent conformational change brings α4 of the receiver domain in one monomer into contact with the beginning of the central domain in the opposite monomer of a dimer (Lee *et al.*, 2000). We propose that upon phosphorylation the receiver domain of the NtrC subgroup may rotate and twist to interact with α7 of the adjacent subunit (Fig. 24). A paper on the activation of NtrC published after completing

this thesis supports this idea (De Carlo *et al.*, 2006). It has been demonstrated for PspF by cryogenic electron microscopy, that two loops (the L1 GAFTGA loop inserted into helix 8 and the L2 loop connecting α9 and the β sheet) are involved in σ^{54} binding, resulting in a transient conformation (Rappas *et al.*, 2005). The interaction of the receiver domain and α7 described above would contribute to the formation of the transient conformation by pulling α7 upwards around the Arg188-D274 axis and thereby decreasing the energy barrier for the activation, which results in the "opening up" of α8 and exposing of the GAFTGA loop to the σ^{54}-RNAP holoenzyme to activate transcription.

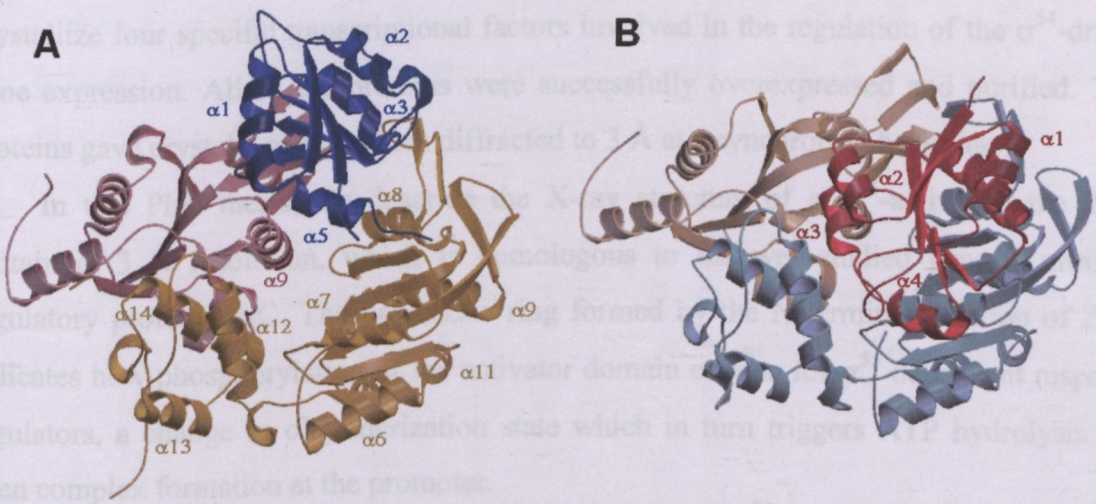


Fig. 24. Crystal structure of the ZraR AAA+ domain to which the N-terminal domain of NtrC has been modelled. (**A**) In the unphosphorylated state, the receiver domain (blue) belonging to molecule A (yellow) interacts with the adjacent molecule B (brown) – however, in this position the formation of a stable hexamer is sterically impossible due to the lack of hydrophobic interactions between α9 and the turn between ß7 and α8 on the other side. (**B**) In the active state, the receiver domain (red) of molecule A rotates and twists to the cleft between the two subdomains of molecule C (cyan).

Summary

Transcription in prokaryotic cells requires the active role of the σ general transcription factor, which has many subclasses, including the alternative σ^{54} protein. The σ^{54}-dependent transcription in bacteria is associated with various stress and growth conditions. Activators of the σ^{54} protein contain a central domain belonging to the AAA+ superfamily of ATPases, members of which function in diverse cellular processes in both prokaryotic and eukaryotic cells, and are often the regulatory component of a two-component signal transduction system.

To gain insight into the activation mechanism of the transcription, we tried to crystallize four specific transcriptional factors involved in the regulation of the σ^{54}-driven gene expression. All of the proteins were successfully overexpressed and purified. Two proteins gave crystals, one of which diffracted to 3 Å at a synchrotron beamline.

In this PhD thesis, we describe the X-ray structure of a σ^{54}-activator, the ZraR protein at 3 Å resolution, which is homologous to the well-studied general nitrogen regulatory protein NtrC. The hexameric ring formed by the N-terminal deletion of ZraR indicates how phosphorylation of the activator domain causes, for σ^{54} dependent response regulators, a change in oligomerization state which in turn triggers ATP hydrolysis and open complex formation at the promoter.

References

1. **Aiba H., Mizuno T. and Mizushima S.** Transfer of phosphoryl group between two regulatory proteins involved in osmoregulatory expression of the ompF and ompC genes in Escherichia coli. ***J. Biol. Chem. 264***, 8563-8567. (1989)
2. **Bordes P., Wigneshweraraj S. R., Schumacher J., Zhang X., Chaney M and Buck M.** The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: identifying a surface that binds sigma 54. ***Proc. Natl. Acad. Sci. USA., 100***, 2278-2283. (2003)
3. **Brissette J. L., Russel M., Weiner L. and Model P.** Phage shock protein, a stress protein of *Escherichia coli*. ***Proc. Natl. Acad. Sci. USA 87***, 862-866. (1990)
4. **Brissette J. L., Weiner L., Ripmaster T. L. and Model P.** Characterization and sequence of the *Escherichia coli* stress-induced *psp* operon. ***J. Mol. Biol. 220***, 35-48. (1991)
5. **Brunger A. T., Adams P. D., Clore G. M., DeLano W. L., Gros P., Grosse-Kunstleve R. W., Jiang J. S., Kuszewski J., Nilges M., Pannu N. S., Read R. J., Rice L. M., Simonson T. and Warren G. L.** Crystallography & NMR system: A new software suite for macromolecular structure determination. ***Acta Crystallogr., D54***, 905-921. (1998)
6. **Chaney M., Grande R., Wigneshweraraj S. R., Cannon W., Casaz P., Gallegos M. T., Schumacher J., Jones S., Elderkin S., Dago A. E., Morett E. and Buck M.** Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminium fluoride: insights into activator mechanochemical action. ***Genes Dev. 15***, 2282-2294. (2001)
7. **Chopra A. K., Peterson J. W. and Prasad R.** Cloning and sequence analysis of hydrogenase regulatory genes (*hyd*HG) from *Salmonella typhimurium*. ***Bioch. Bioph. Acta 1129***, 115-118. (1991)
8. **Collaborative Computational Project, Number 4.** ***Acta Cryst. D50***, 760-763. (1994)
9. **Cserzo M., Wallin E., Simon I., Heijne G. von and Elofsson A.** Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. ***Protein Eng. 10***, 673-676. (1997)

10. **De Carlo S., Chen B., Hoover T. R., Kondrashkina E., Nogales E. and Nixon B. T.** The structural basis for regulated assembly and function of the transcriptional activator NtrC. ***Genes Dev. 20(11),*** 1485-1495. (2006)
11. **Doucleff M., Chen B., Maris A. E., Wemmer D. E., Kondrashkina E. and Nixon B. T.** Negative regulation of AAA+ ATPase assembly by two component receiver domains: a transcription activation mechanism that is conserved in mesophilic and extremely hyperthermophilic bacteria. ***J. Mol. Biol. 353(2)***, 242-255. (2005)
12. **Drummond M. H., Conteras A. and Mitchenall L. A.** The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of *Klebsiella pneumoniae*. ***Mol. Microbiol. 4***, 29-37. (1990)
13. **Dworkin J., Jovanovic G. and Model P.** The PspA protein of Escherichia coli is a negative regulator of σ^{54}-dependent transcription. ***J. Bacteriol. 182***, 311-319. (2000)
14. **Elderkin S., Jones S., Schumacher J., Studholme D. and Buck. M.** Mechanism of action of the *Escherichia coli* phage shock protein PspA in repression of the AAA family transcription factor PspF. ***J. Mol. Biol. 320***, 23-37. (2002)
15. **Fiedler U. and Weiss V.** A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules. ***EMBO J. 14***, 3696-3705. (1995)
16. **Flashner Y., Weiss D. S., Keener J. and Kustu S.** Constitutive forms of the enhancer-binding protein NtrC: evidence that essential oligomerization determinants lie in the central activation domain. ***J. Mol. Biol. 249***, 700-713. (1995)
17. **Graaff R. A. G. de, Hilge M., Plas J. L. van der and Abrahams J. P.** Matrix methods for solving protein substructures of chlorine and sulfur from anomalous data. ***Acta Crystallogr. D57***, 1857-1862. (2001)
18. **Hankamer B. D., Elderkin S. L., Buck M. and Nield J.** Organization of the AAA(+) adaptor protein PspA is an oligomeric ring. ***J. Biol. Chem. 279***, 8862-8866. (2004)
19. **Hanson P. I. and Whiteheart S. W.** AAA+ proteins: have engine, will work. ***Nat. Rev. Mol. Cell Biol. 6***: 519-529. (2005) Review.
20. **Hilgenfeld R.** Regulatory GTPases. ***Curr. Opin. Struct. Biol. 5***, 810-817. (1995) Review.
21. **Hishida T., Han Y. W., Fujimoto S., Iwasaki H. and Shinagawa H.** Direct evidence that a conserved arginine in RuvB AAA+ ATPase acts as an allosteric

effector for the ATPase activity of the adjacent subunit in a hexamer. ***Proc. Natl. Acad. Sci. USA. 101***, 9573-9577. (2004)

22. **Holm L. and Sander C.** Protein structure comparison by alignment of distance matrices. ***J. Mol. Biol. 233***, 123-138. (1993)
23. **Hoover T. R., Santero E., Porter S. and Kustu S.** The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. ***Cell 63***, 11-22. (1990)
24. **Hwang I., Thorgeirsson T., Lee J., Kustu S. and Shin Y. K.** Physical evidence for a phosphorylation-dependent conformational change in the enhancer-binding protein NtrC. ***Proc. Natl. Acad. Sci. USA. 96***, 4880-4885. (1999)
25. **Iyer L. M., Leipe D. D., Koonin E. V. and Aravind L.** Evolutionary history and higher order classification of AAA+ ATPases. ***J Struct Biol. 146(1-2)***, 11-31. (2004)
26. **Jahn R., Lang T. and Sudhof T.C.** Membrane fusion. ***Cell. 112(4)***, 519-33. (2003) Review.
27. **Jones T. A., Zou J. Y., Cowan S. W. and Kjeldgaard M.** Improved methods for building protein models in electron density maps and the location of errors in these models. ***Acta Crystallogr. A47***, 110-119. (1991)
28. **Jovanovic G. and Model P.** PspF and IHF bind co-operatively in psp promoter regulatory regions of *Escherichia coli.* ***Mol. Micobiol. 25***, 473-481. (1997)
29. **Jovanovic G., Weiner L. and Model P.** Identification, nucleotide sequence, and characterization of PspF, the transcriptional activator of the *Escherichia coli* stress-induced *psp* operon. ***J. Bacteriol. 178***, 1936-1945. (1996)
30. **Keener J. and Kustu S.** Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. ***Proc. Natl. Acad. Sci. USA. 85***, 4976-4980. (1988)
31. **Kern D., Volkman B. F., Luginbuhl P., Nohaile M. J., Kustu S. and Wemmer D. E.** Structure of a transiently phosphorylated switch in bacterial signal transduction. ***Nature 402***, 894-898. (1999)
32. **Kleerebezem M., Crielaard W. and Tommassen J.** Involvement of stress protein PspA (phage shock protein A) of *Escherichia coli* in maintenance of the protonmotive force under stress conditions. ***EMBO J. 15***, 162-171. (1996)
33. **Kleerebezem M. and Tommassen J.** Expression of the pspA gene stimulates efficient protein export in *Escherichia coli*. ***Mol. Microbiol. 7***, 947-956. (1993)

34. **Klose K. E., North A. K., Stedman K. M. and Kustu S.** The major dimerization determinants of the nitrogen regulatory protein NTRC from enteric bacteria lie in its carboxy-terminal domain. ***J. Mol. Biol. 241***, 233-245. (1994)

35. **Kraulis P. J.** MOLSCRIPT: a program to produce detailed and schematic plots of protein structure. ***J. Appl. Cryst. 24***, 946-950. (1991)

36. **La Fortelle E. de and Bricogne G.** Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. ***Methods Enzymol. 276***, 472-494. (1997)

37. **Laskowski R. A., MacArthur M. W., Moss D. S. and Thornton J. M.** PROCHECK - a program to check the stereochemical quality of protein structures. ***J. Appl. Cryst. 26***, 283-291. (1993)

38. **Lee S. Y., De La Torre A., Yan D., Kustu S., Nixon B. T. and Wemmer D. E.** Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. ***Genes Dev. 17***, 2552-2563. (2003)

39. **Lee J., Owens J. T., Hwang I., Meares C. and Kustu S.** Phosphorylation-induced signal propagation in the response regulator ntrC. ***J. Bacteriol. 182***, 5188-5195. (2000)

40. **Lenzen C. U., Steinmann D., Whiteheart S. W. and Weiss W. I.** Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive protein. ***Cell 94***, 525-536. (1998)

41. **Leonhartsberger S., Huber A., Lottspeich F. and Böck A.** The *hydH/G* genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system. ***J. Mol. Biol. 307***, 93-105. (2001)

42. **Lew C. M. and Gralla J. D.** New roles for conserved regions within a sigma 54-dependent enhancer-binding protein. ***J. Biol. Chem. 277***, 41517-41524. (2002)

43. **Lloyd L. J., Jones S. E., Jovanovic G., Gyaneshwar P., Rolfe M. D., Thompson A., Hinton J. C. and Buck M.** Identification of a new member of the phage shock protein response in Escherichia coli, the phage shock protein G (PspG). ***J. Biol. Chem. 279***, 55707-55714. (2004)

44. **Lois A. F., Weinstein M., Ditta G. S. and Helinski D. R.** Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of Rhizobium meliloti are coordinately regulated by oxygen. ***J. Biol. Chem. 268***, 4370-4375. (1993)

45. **Loomis W. F., Shaulsky G. and Wang N.** Histidine kinases in signal transduction pathways of eukaryotes. ***J. Cell Sci. 110***, 1141-1145. (1997) Review.

46. **Lu C. D. and Abdelal A. T.** Role of ArgR in activation of the ast operon, encoding enzymes of the arginine succinyltransferase pathway in *Salmonella typhimurium*. ***J. Bacteriol. 191***, 1934-1938. (1999)

47. **MacFarlane S. A. and Merrick M.** The nucleotide sequence of the nitrogen regulation gene ntrB and the glnA-ntrBC intergenic region of *Klebsiella pneumoniae*. ***Nucleic Acids Res. 13***, 7591-7606. (1985)

48. **Maeda T., Wurgler-Murphy S. M. and Saito H.** A two-component system that regulates an osmosensing MAP kinase cascade in yeast. ***Nature 369***, 242-245. (1994)

49. **Merritt E. A. and Bacon D. J.** Raster3D Photorealistic Molecular Graphics. ***Methods Enzymol. 277***, 505-524. (1997)

50. **Mettke I., Fiedler U. and Weiss V.** Mechanism of activation of a response regulator: interaction of NtrC-P dimers induces ATPase activity. ***J. Bacteriol. 177***, 5056-5061. (1995)

51. **Meyer M. G., Park S., Zeringue L., Staley M., McKinstry M., Kaufman R. I., Zhang H., Yan D., Yennawar N., Yennawar H., Farber G. K. and Nixon B. T.** A dimeric two-component receiver domain inhibits the sigma54-dependent ATPase in DctD. ***FASEB J. 15***, 1326-1328. (2001)

52. **Mizuno, T.** Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. ***DNA Res. 4***, 161-168. (1997)

53. **Mogk A., Schlieker C., Strub C., Rist W., Weibezahn J. and Bukau B.** Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. ***J. Biol. Chem. 278***, 17615-17624. (2003)

54. **Murshudov G. N., Vagin A. A. and Dodson E. J.** Refinement of macromolecular structures by the maximum-likelihood method. ***Acta Crystallogr. D53***, 240-255. (1997)

55. **Neuwald A. F., Aravind L., Spouge J. L. and Koonin E. V.** AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. ***Genome Res. 9***, 27-43. (1999)

56. **Nohaile M., Kern D., Wemmer D., Stedman K. and Kustu S.** Structural and functional analyses of activating amino acid substitutions in the receiver domain of NtrC: evidence for an activating surface. ***J. Mol. Biol. 273***, 299-316. (1997)

57. **North A. K., Weiss D. S., Suzuki H., Flashner Y. and Kustu S.** Repressor forms of the enhancer-binding protein NtrC: some fail in coupling ATP hydrolysis to open complex formation by sigma 54-holoenzyme. ***J. Mol. Biol. 260***, 317-331. (1996)

58. **Ohnishi T., Hishida T., Harada Y., Iwasaki H. and Shinagawa H.** Structure-Function Analysis of the Three Domains of RuvB DNA Motor Protein. ***J. Biol. Chem. 280***, 30504-30510. (2005)

59. **Osuna J., Soberon X. and Morett E.** A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. ***Protein Sci. 6***, 543-555. (1997)

60. **Otwinowski Z. and Minor W.** Processing of X-ray Diffraction Data Collected in Oscillation Mode. ***Methods Enzymol. 276***, 307-326. (1997)

61. **Park S., Meyer M., Jones A. D., Yennawar H. P., Yennawar N. H. and Nixon B. T.** Two-component signaling in the AAA+ ATPase DctD: binding Mg2+ and BeF3-selects between alternate dimeric states of the receiver domain. ***FASEB J. 16***: 1964-1966. (2002)

62. **Patel S. and Latterich M.** The AAA team: related ATPases with diverse functions. ***Trends Cell Biol. 8***, 65-71. (1998) Review.

63. **Peck H. D. and Gest H.** Formic dehydrogenase and the hydrogenlyase enzyme complex in coli-aerogenes bacteria. ***J. Bacteriol. 73***, 706-721. (1957)

64. **Pelton J. G., Kustu S. and Wemmer D. E.** Solution structure of the DNA-binding domain of NtrC with three alanine substitutions. ***J. Mol. Biol. 292***, 1095-1110. (1999)

65. **Perez-Martin J. and de Lorenzo V.** ATP binding to the sigma 54-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS DNA. ***Cell 86***, 331-339. (1996)

66. **Pickhart C. M. and Cohen R. E.** Proteasomes and their kin: proteases in the machine age. ***Nat. Rev. Mol. Cell Biol. 5***, 177-87. (2004). Review.

67. **Popham D. L., Szeto D., Keener J. and Kustu S.** Function of a bacterial activator protein that binds to transcriptional enhancers. ***Science 243***, 629-635. (1989)

68. **Putnam C. D., Clancy S. B., Tsuruta H., Gonzalez S., Wetmur J. G. and Tainer J. A.** Structure and mechanism of the RuvB Holliday junction branch migration motor. ***J. Mol. Biol. 311***, 297-310. (2001)

69. **Rappas M., Schumacher J., Beuron F., Niwa H., Bordes P., Wigneshweraraj S., Keetch C. A., Robinson C. V., Buck M. and Zhang X.** Structural insights into the activity of enhancer-binding proteins. ***Science 307***, 1972-1975. (2005)

70. **Rappas M., Schumacher J., Niwa H., Buck M. and Zhang X.** Structural basis of the nucleotide driven conformational changes in the AAA+ domain of transcription activator PspF. ***J. Mol. Biol. 357***, 481-492. (2006)

71. **Reitzer L.** Nitrogen assimilation and global regulation in Escherichia coli. ***Annu. Rev. Microbiol. 57***, 155-176. (2003) Review.

72. **Reitzer L. J. and Magasanik B.** Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. ***Cell 45***, 785-792. (1986)

73. **Reitzer L. and Schneider B.** Metabolic context and possible physiological themes of σ^{54}-dependent genes in *Escherichia coli*. ***Microbiol. Mol. Biol. Rev. 65***, 422-444. (2001) Review.

74. **Rippe K., Mucke N. and Schulz A.** Association states of the transcription activator protein NtrC from *E. coli* determined by analytical ultracentrifugation. ***J. Mol. Biol. 278***, 915-933. (1998)

75. **Rombel I., North A., Hwang I., Wyman C. and Kustu S.** The bacterial enhancer-binding protein NtrC as a molecular machine. ***Cold Spring Harb. Symp. Quant. Biol. 63***, 157-166. (1998) Review.

76. **Rombel I., Peters-Wendisch P., Mesecar A., Thorgeirsson T., Shin Y. K. and Kustu S.** MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. ***J. Bacteriol. 181***, 4628-4638. (1999)

77. **Rouiller I., DeLaBarre B., May A. P., Weis W. I., Brunger A. T., Milligan R. A. and Wilson-Kubalek E. M.** Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. ***Nat. Struct. Biol. 9***, 950-957. (2002)

78. **Saraste M., Sibbald P. R. and Wittinghofer A.** The P-loop–a common motif in ATP- and GTP-binding proteins. ***Trends Biochem. Sci. 15***, 430-434. (1990) Review.

79. **Schneider T. R. and Sheldrick G. M.** Substructure solution with SHELXD. ***Acta Cryst. D58***, 1772-1779. (2002)

80. **Schumacher J., Zhang X., Jones S., Bordes P. and Buck M.** ATP-dependent transcriptional activation by bacterial PspF AAA+ protein. ***J. Mol. Biol. 14***; 863-875. (2004)

81. **Shingler V.** Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism. ***Mol. Microbiol. 19***, 409-416. (1996) Review.

82. **Stock A. M., Robinson V. L. and Goudreau P. N.** Two-component signal transduction. ***Annu. Rev. Biochem. 69***, 183-215. (2000) Review.

83. **Stoker K., Reijnders W. M., Oltmann F. and Stouthamer A. H.** Initial cloning and sequencing of *hyd*HG, an operon homologous to *ntr*BC and regulating the labile hydrogenase activity in *Escherichia coli* K12. ***J. Bacteriol. 171***, 4448-4456. (1989)

84. **Studholme D. J., Wigneshweraraj S. R., Gallegos M. T. and Buck M.** Functionality of purified σ^N (σ^{54}) and a NifA-like protein from the hyperthermophile Auqifex aeolicus. ***J. Bacteriol. 182***, 1616-1623. (2000)

85. **Su W., Porter S., Kustu S. and Echols H.** DNA-looping and enhancer activity: Association between DNA-bound NTRC activator and RNA polymerase at the bacterial *glnA* promoter. ***Proc. Natl. Acad. Sci. 87***, 5504-5508. (1990)

86. **Surette M. G., Levit M., Liu Y., Lukat G., Ninfa E. G., Ninfa A. and Stock J. B.** Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis. ***J. Biol. Chem. 271***, 939-945. (1996)

87. **Terwilliger T. C.** Automated structure solution, density modification and model building. ***Acta Crystallogr. D58***, 1937-1940. (2002)

88. **Venien-Bryan C., Schertler G. F., Thouvenin E. and Courty S.** Projection structure of a transcriptional regulator, HupR, determined by electron cryo-microscopy. ***J. Mol. Biol. 296***, 863-871. (2000)

89. **Walker J. E., Saraste M., Runswick M. J. and Gay M. J.** Distantly related sequences in the α- and β- subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. ***EMBO J. 1***, 945-951. (1982)

90. **Wang J., Song J. J., Seong I. S., Franklin M. C., Kamtekar S., Eom S. H. and Chung C. H.** Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU. ***Structure 9***, 1107-1116. (2001)

91. **Weeks C. M. and Miller R.** The design and implementation of SnB v2.0. ***J. Appl. Cryst. 32***, 120-124. (1999)

92. **Weiner L., Brissette J. L., Ramani N. and Model P.** Analysis of proteins and cis-acting elements regulating the stress-induced phage-shock protein operon. ***Nucl. Acids Res. 23***, 2030-2036. (1995)

93. **Weiner L. and Model P.** Role of an Escherichia coli stress-response operon in stationary-phase survival. ***Proc. Natl. Acad. Sci. USA 91***, 2191-2195. (1994)

94. **Wyman C., Rombel I., North A. K., Bustamante C. and Kustu S.** Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. ***Science 275***, 1658-1661. (1997)

95. **Yan D. and Kustu S.** "Switch I" mutant forms of the bacterial enhancer-binding protein NtrC that perturb the response to DNA. ***Proc. Natl. Acad. Sci. USA. 96***, 13142-13146. (1999)

96. **Yu R. C., Hanson P. I., John R. and Brünger A. T.** Structure of the ATP-dependent oligomerization domain of the N-ethylmaleimide-sensitive factor complexed with ATP. ***Nat. Struct. Biol. 5***, 803-811. (1998)

97. **Yu R. C., Jahn R. and Brunger A. T.** NSF N-terminal domain crystal structure: models of NSF function. ***Mol. Cell 4***, 97-107. (1999)

Appendix